AD_____
(Leave blank)

Award Number: W81XWH-07-1-0619

TITLE: Genetic and Environmental Pathways in Type 1 Diabetes Complications

PRINCIPAL INVESTIGATOR: Massimo Trucco, M.D.

CONTRACTING ORGANIZATION: Children's Hospital of Pittsburgh
Pittsburgh, PA 15213

REPORT DATE: September 2008

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: (Check one)

X Approved for public release; distribution unlimited

Distribution limited to U.S. Government agencies only; report contains proprietary information

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
26-09-2008	Annual	27 Aug 2007-26 Aug 2008

4. TITLE AND SUBTITLE
Genetic and Environmental Pathways in Type 1 Diabetes Complications

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-07-1-0619

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Massimo Trucco, M.D.

Email: mnt@pitt.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Children's Hospital of Pittsburgh of
UPMC Health System
3705 Fifth Avenue at DeSoto Street
Pittsburgh, PA 15213

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research
and Materiel Command
Fort Detrick, MD 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Genetic factors contribute to risk for developing nephropathy in patients with Type 1 Diabetes (T1D). Cigarette smoking is deleterious to kidney function and is a risk factor for Diabetic-Nephropathy (DN) as well as End Stage Renal Disease (ESRD) in patients with T1D. The proposed study investigates how environmental exposure(s) (e.g., smoking) and genetic variants interact to amplify risk for T1DN and substantially increase incidence of ESRD. The specific aims are: 1) Identify genetic variants conferring risk to T1DN by performing a staged follow-up of our initial Genome-Wide Association Scan (GWAS) results; 2) Ensure that SNPs identified by Aim 1 affect risk of T1DN, as opposed to risk for T1D; 3) Identify genetic variants that interact with smoking status in conferring risk for T1DN; 4) Confirm results obtained during Aims 1-3 using an independent cohort of case and control participants. The relevance of the study to public health is that 16 million people in the US have diabetes with 800,000 new cases diagnosed each year. Diabetic complications threatening vision, kidney, and nerve function affect most diabetic patients. Improved prediction of risk for developing diabetes and diabetic complications among active duty members of the military, their families and retired military personnel will potentially allow focused preventative treatment of at-risk individuals, providing significant healthcare savings and improved patient well being.

15. SUBJECT TERMS
End Stage Renal Disease; Genetic Association; Genome Scanning; Nephropathy; Type 1 Diabetes

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
U	U	U	UU	95	**19b. TELEPHONE NUMBER** *(include area code)*

Children's Hospital of Pittsburgh
W81XWH-07-1-0619
Annual Report (08/27/2006 – 08/26/2008)
Table of Contents

ASHI Quarterly 31:50-52, 2007.

Methods Mol Biol 373:25-38, 2007.

Pediatr Diabetes 8:307-322, 2007

Am J Hum Genet 82:453-463, 2008.

INTRODUCTION:

Type 1 Diabetes (T1D) is associated with increased risk of T1D-Nephropathy (T1DN) and is usually accompanied by other diabetic-related complications such as retinopathy, neuropathy, blood pressure elevation, and high risk of cardiovascular morbidity and mortality. Sixteen million people in the US have diabetes with 800,000 new cases diagnosed each year. Diabetic complications affect most diabetic patients. Diabetes occurs in men, women, children and the elderly. African, Hispanic, Native and Asian Americans are particularly susceptible to its most severe complications. An estimated 20% to 40% of T1D patients will develop diabetic nephropathy, clinically first evidenced by microalbuminuria, during their lifetime. If untreated nearly all T1D patients experiencing microalbuminuria will progress to overt nephropathy, evidenced by macroalbuminuria, and culminating in T1D-End Stage Renal Disease (T1D-ESRD). Improved prediction of risk for developing diabetes and diabetic complications among active duty members of the military, their families and retired military personnel will potentially allow focused preventative treatment of at-risk individuals, providing significant healthcare savings and improved patient well being.

BODY:

Our first quarterly scientific progress report (08/27/07 – 11/30/07) detailed the following steps forward in reaching the aims of our study.

The proposed study will investigate how environmental exposure(s) and genetic variants interact to amplify risk for Type 1 Diabetic (T1D) nephropathy (T1DN) and substantially increase incidence of end-stage renal disease (T1D-ESRD). The work that will be performed during the proposed 2 year study will exploit the results of our previously funded application to the DOD entitled "*New Advanced Technology to Improve Prediction of Type 1 Diabetes and Its Complications*", the aims of which are nearing completion. The previous project produced 3 principal results: 1) Transmission/Disequilibrium Testing (TDT) analysis using a cohort of 92 case family trios for the association of 121 candidate genes with T1DN, identifying significant association between T1DN and 4 of the loci tested; 2) Recruitment of an independent cohort of 2,881 participants, consisting of 1,174 case/control singletons and 569 case/control family trios; and 3) a genome-wide association scan (GWAS) of 197 singleton cases and 197 controls all of whom were non-smokers.

The currently proposed work follows up on the significant genetic signals identified during our previous project.

To accomplish our aims we originally proposed the following tasks:

Task 1. Using our available cohort of 257 T1DN (case) family trios perform an additional series of GWAS using the 500,000 SNP typing microarray in order to confirm the strongest genetic signals associated with diabetic kidney disease.

Task 2. Recruit an additional 200 participants with T1DN who have never smoked cigarettes in order to increase the total number of singleton case participants to greater than 280.

Task 3. Design customized microarrays in order to examine the significant genetic markers using an independent cohort of singleton cases and controls matched for exposure to cigarette smoking. The available cohort for this stage of the project is: Non-Smokers, 86 case and 201 control; Smokers, 295 case and 198 control.

Task 4. Perform SNP genotyping of the most significant genetic markers using an available control cohort consisting of 312 family trios matched for exposure to cigarette smoking: Non-Smokers, 216; Smokers, 96.

Task 5. In order to complete the analysis of the genetic and environmental exposure data for association with T1DN we will compare the genetic marker frequency (i.e., SNPs, Copy Number Variation, Haplotype, Gene-Gene, and Gene-Environment interactions) with occurrence of the disease.

This quarterly report provides a perfect opportunity to present an update of our progress regarding the T1D genome-wide scan (GWAS). The technical and statistical solutions successfully used to confirm the genetic regions most associated T1D will then be used to investigate T1DN associations. A cohort of 422 T1D and 2,144 non-T1D participants were used to investigate whether it would be possible to identify novel SNPs for association with the phenotype (Table 1). The T1D participants

were recruited from the Children's Hospital of Pittsburgh (N=28) and obtained from the GoKinD study (N=394). Non-T1D participants were provided by the KORA and POPGEN studies and were received as computer files of genotype information. All participants, whether collected in Pittsburgh or provided by KORA/POPGEN, were genotyped using the Affymetrix 500K GeneChip. This is the same SNP typing array that was used by the Wellcome Trust Case Control Consortium (WTCCC). We are in the process of applying for access to the WTCCC dataset and anticipate having a positive reply before February 2008.

Table 1. Characteristics of Case and Control Participants.

	Case Participants		Control Participants	
	CHP	GoKinD	KORA	POPGEN
Demographic Characteristics:				
Number of Singletons	28	394	1644	500
European American (%)	100%	100%	----	----
German Resident (%)	----	----	100%	100%
Male Gender (%)	50%	46.7%	49.5%	51.8%
History of Diabetes:				
Type 1 Diabetes (%)	100%	100%	----	----
Mean Age at T1D Diagnosis (yr)	12.7(7.9)	12.2(7.1)	----	----

Values in parentheses indicate standard deviation from the mean.

As expected, we observed a large signal corresponding to the HLA region that exceeded genome-wide significance (Figure 1). Other notable signals were observed but fell short of the Bonferroni correction for the threshold of significance (p-value less than $4x10^{-7}$). Suggestive signals observed for several previously identified T1D loci are summarized in Table 2. For example, the *PHTF1/PTPN22* loci ($p=1x10^{-5}$), *CTLA4* ($p=3x10^{-4}$), and *IL2RA* ($p=1x10^{-3}$) are commonly considered as confirmed susceptibility loci associated with T1D. Other suggested T1D susceptibility loci (e.g., *ITPR3*, *NOD2*, and *PDX1*) were also identified by the GWAS. In contrast, the confirmed loci *CD25*, *IFIH1*, and *INS* were not readily observed as nearby SNPs did not exceed a p-value of 0.001. This was possibly the result of the Affymetrix GeneChip not including a sufficient number of SNPs in linkage disequilibrium with the implicated alleles.

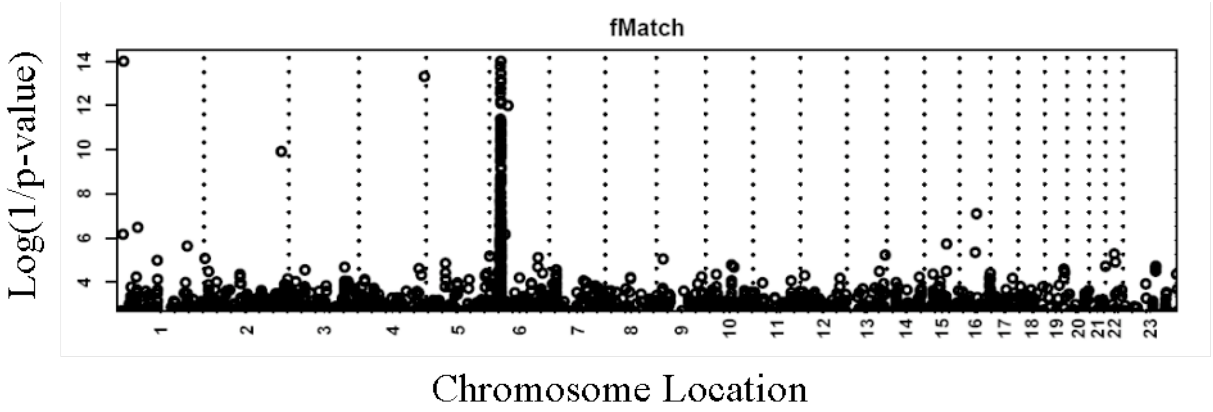

Figure 1. Summary of the GWAS data obtained for T1D. The x-axis is chromosome and location while the y-axis is the -Log of the p-value. As expected the HLA region on chromosome 6p attained genome-wide significance. All other p-values exceeding $1x10^{-7}$ have been removed from consideration due to poor quality genotyping data. The complete list of GWAS data is included in the Appendix.

Table 2. Confirmed and suggested T1D loci identified during the T1D GWAS.

Chr	Locus	Location	dbSNP_ID	Dominant	Additive	Recessive
1	PHTF1	114,045,700	rs1230649	1.2E-02	**7.2E-04**	1.0E-01
1	PHTF1	114,105,331	rs6679677	**3.3E-05**	**1.0E-05**	4.4E-02
1	PTPN22	114,216,891	rs2488457	**3.8E-03**	**7.1E-04**	1.2E-01
2	CTLA4	204,428,384	rs11571292	2.5E-02	**4.0E-03**	**2.6E-04**
2	CTLA4	204,449,111	rs231726	**7.3E-03**	**3.4E-04**	**5.2E-04**
2	CTLA4	204,449,411	rs1427676	**6.5E-03**	**4.5E-04**	**6.0E-04**
6	ITPR3	33,650,501	rs9296095	**2.2E-04**	**3.1E-03**	1.0E-02
6	ITPR3	33,682,987	rs444697	**2.5E-04**	**3.9E-05**	**5.1E-04**
10	IL2RA	6,132,099	rs10905669	6.2E-02	**1.0E-03**	3.9E-02
13	PDX1	27,374,640	rs9554196	**6.5E-04**	1.8E-02	**9.5E-03**
13	PDX1	27,377,980	rs9319399	**3.8E-04**	1.6E-02	**7.5E-03**
13	PDX1	27,378,356	rs9551419	**7.4E-04**	2.2E-02	**9.5E-03**
16	NOD2	49,281,588	rs9933594	**9.1E-04**	1.3E-02	**6.3E-03**

Table 3 lists select data obtained from the GWAS. Of the loci identified, there were 12 that are currently listed by T1DBase (http://t1dbase.org) as candidates for T1D. Moreover, there are a number of novel loci that may be interesting to evaluate during a staged investigation of the phenotype. For example, the GWAS identified loci involved in insulin signaling (*CRK* and *IRS2*), cytokine signaling (*IL6ST*, *IRF4*, *TRAF3Ip1*) and potassium ion transport (*KCNH7* and *KCNQ1*). Moreover, the GWAS identified a subunit of mitochondrial complex 1 (*NDUFA10*) and the apoptosis regulatory factor *FKSG2* that may influence superoxide production and subsequent susceptibility to programmed cell death.

Table 3. Select loci observed during the T1D GWAS.

Chr	Locus	Location	dbSNP_ID	Dominant	Additive	Recessive
T1DBase Candidate Loci:						
1	IL12RB2	67,612,161	rs3828069	**5.7E-03**	**9.3E-03**	**9.8E-04**
1	PHTF1/PTPN22	114,105,331	rs6679677	**3.3E-05**	**1.0E-05**	4.4E-02
2	CTLA4	204,428,384	rs11571292	2.5E-02	**4.0E-03**	**2.6E-04**
3	SENP7	102,504,337	rs16843699	**8.5E-04**	**7.5E-04**	6.8E-02
6	ITPR3	33,682,987	rs444697	**2.5E-04**	**3.9E-05**	**5.1E-04**
10	IL2RA	6,132,099	rs10905669	6.2E-02	**1.0E-03**	3.9E-02
11	LRP5	67,813,346	rs11606508	**2.2E-03**	**3.1E-03**	**9.4E-04**
12	CD4	6,775,811	rs1075836	6.0E-02	**2.7E-04**	5.9E-02
12	TRAFD1	111,002,025	rs1980364	**9.6E-04**	**9.4E-03**	**5.3E-04**
12	PTPN11	111,462,160	rs10850053	**6.7E-04**	**9.2E-03**	**5.5E-04**
13	PDX1	27,377,980	rs9319399	**3.8E-04**	1.6E-02	**7.5E-03**
16	NOD2	49,281,588	rs9933594	**9.1E-04**	1.3E-02	**6.3E-03**
Novel T1D Candidate Loci:						
2	KCNH7	163,442,376	rs10203602	**1.1E-03**	**2.4E-04**	**3.7E-03**
2	TRAF3IP1	238,928,000	rs821804	**1.7E-05**	**7.1E-05**	**4.4E-05**
2	NDUFA10	240,555,969	rs4854067	**1.6E-03**	2.5E-02	**1.2E-04**
5	IL6ST	55,388,527	rs10042443	**2.2E-04**	**1.4E-05**	**1.4E-04**
6	IRF4	374,457	rs950286	**5.2E-05**	**6.7E-06**	**4.8E-03**
8	FKSG2	36,899,367	rs1010669	6.3E-02	1.6E-02	**3.9E-05**
8	EYA1	72,362,997	rs4738123	**5.0E-03**	**6.2E-05**	**8.7E-04**
9	IL6RL1	91,744,788	rs7864499	**2.0E-04**	**5.1E-04**	**2.6E-03**
11	KCNQ1	2,781,804	rs163166	**4.4E-05**	1.7E-02	**3.3E-03**
13	IRS2	109,307,346	rs1041637	**1.6E-03**	**6.0E-06**	**1.8E-04**
15	SEMA6D	45,680,428	rs16959754	**8.7E-06**	**9.1E-05**	**6.3E-06**
17	CRK	1,278,556	rs8070640	**2.7E-05**	**9.4E-04**	**5.7E-05**

The complete list of SNPs implicated during the GWAS has been included in the Appendix. That file contains the chromosome location of each SNP as well as the p-value associated with the 3 major modes of inheritance (i.e., additive, dominant, and recessive). The Appendix also summarizes the quality control analysis for those SNPs with p-values less than 0.0001 and indicates which SNPs overlap T1D candidate loci that have been listed by T1DBase. This is the principal file that is being used to determine which signals obtained during the GWAS for T1D should be studied during Stage 2 analysis of the T1D results.

There are expected to be 14,047 participants available for the T1D project (Table 4). Our repository is anticipated to contain a modest number of T1D family trios (N=719), T1D singletons (N=4,246), and non-T1D control singletons (N=7,644). Table 4 summarizes the available cohorts and details the source of each participant, and whether material is provided as DNA or as a computer file of previously determined genotypes. When genotypes are provided they were obtained using the same Affymetrix 500K GeneChip that was used during our GWAS for T1D.

Table 4. Summary of cohort available and anticipated for T1D analysis.

	Family Trios	Case and Control Singletons		
Source	T1D Family Trios	Case	Control	Comment
EMIL	150	300	2500	DNA available
GoKinD	569	1746		DNA available
KORA			1644	GeneChip data available
POPGEN			500	GeneChip data available
Pittsburgh		200		DNA available
WTCCC		2000	3000	GeneChip data available

At this stage of the project we need to prioritize the SNPs identified during the GWAS and then evaluate suggestive SNPs using the T1D family trios. The University of Pittsburgh Genomics Core Laboratory has recently acquired the Sequenom SNP typing platform. The Sequenom system provides sufficient throughput and genotyping accuracy for projects such as ours. The fee for using this system has not yet been finalized but is anticipated to be about 10 cents per SNP. This is about one-fifth the fee associated with genotyping using TaqMan methodology. It is anticipated that when working with T1D family trios we will genotype 2,157 participants (N=3x719) for about 45 SNPs at a cost of roughly $9,700. Choosing the 45 SNPs to genotype is the aim that needs to be completed in the near future.

The next phase of analysis will have 2 goals. The first being to measure the p-values for confirmed and suggested T1D loci and the second being to test whether the signals associated with novel T1D loci are reproducible.

12. Statement of Plans and Milestones for the Next Quarter

There are 2 goals (Goal 1 and 2) that we anticipate completing and 1 goal (Goal 3) that will be initiated during the next research period.

Goal 1. Increase the number of participants. **Milestone 1A.** Apply to the WTCCC for access to the T1D and control cohorts. **Milestone 1B.** Continue to recruit T1D case trios and singletons from Pittsburgh.

Our application to the WTCCC for access to their cohort of T1D (N=2,000) and non-T1D control (N=3,000) singletons has been finalized. The project title is "Detection of Susceptibility Loci in Multifactorial Disease". The principal aim is to combine our original GWAS results with the GWAS performed by WTCCC followed by statistical genetic analysis. It is anticipated that the outcome will be an analysis with increased power to detect true signals for association with T1D risk. We will also continue to recruit T1D participants through our collaboration with the University of Pittsburgh Transplantation Institute. Samples collected from Pittsburgh will

be used to increase the size of our DNA repository enabling increased power of the genetic study to detect true association signals.

Goal 2. Determine high priority SNPs to study. **Milestone 2A.** Identify confirmed and/or suggestive T1D risk loci using our original GWAS. **Milestone 2B.** Identify novel loci to evaluate for T1D risk.

Data obtained during our original GWAS will be compared with the list of confirmed and suggested T1D risk loci in order to identify where the different datasets overlap. Moreover, novel SNPs identified during the original GWAS will be complied based upon the following criteria: i) SNPs with p-values less than 0.0001; ii) SNPs with p-values less than 0.001 occurring within 50,000bp of a known locus; and iii) Clusters of 3 or more SNPs with p-vlaues less than 0.001 occurring within 50,000bp of one another. The priority SNPs will be used to compile a list that will direct subsequent laboratory analyses.

Goal 3. Initiate genotyping of priority SNPs. **Milestone 3A.** Use the cohort of T1D family trios to evaluate confirmed SNPs for T1D risk. **Milestone 3B.** Use the cohort of T1D family trios to evaluate novel SNPs for T1D risk.

Our DNA repository includes a collection of T1D family trios (N=719) that can be used to evaluate the SNPs identified during Goal 2. SNPs will be chosen from the list of confirmed and suggestive T1D risk as well as from the list of novel genetic variants. The genotype of each SNP will be determined among the various participants and the results will be evaluated by transmission disequilibrium testing to determine association with T1D.

In our second quarterly scientific progress report (12/01/07 – 02/29/08, we presented the following data:

Type 1 Diabetes Genome-Wide Association Study Update:
The previous quarterly report summarized our progress regarding the genome-wide association (GWA) scan for Type 1 Diabetes (T1D) susceptibility alleles. We have reported the genotype results from a cohort of 422 T1D (case) and 2,144 non-T1D (control) participants using 322,347 SNPs. The anthropometric properties of the case and control cohort has been described in the previously quarterly report but is reproduced below for convenience.

Table 1. Characteristics of Case and Control Participants.

	Case Participants		Control Participants	
	CHP	GoKinD	KORA	POPGEN
Demographic Characteristics:				
Number of Singletons	28	394	1644	500
European American (%)	100%	100%	----	----
German Resident (%)	----	----	100%	100%
Male Gender (%)	50%	46.7%	49.5%	51.8%
History of Diabetes:				
Type 1 Diabetes (%)	100%	100%	----	----
Mean Age at T1D Diagnosis (yr)	12.7(7.9)	12.2(7.1)	----	----

Values in parentheses indicate standard deviation from the mean.

As outlined in the following sections, the research goals for the first quarter of 2008 were 3-fold: 1) Increase the number of participants; 2) Determine high priority SNPs to study; and 3) Initiate genotyping of high priority SNPs. Our progress for each goal is addressed below.

Goal 1. *Increase the number of participants.* ***Milestone 1A.*** *Apply to the Wellcome Trust Case Control Consortium (WTCCC) for access to the T1D and control cohorts.* ***Milestone 1B.*** *Continue to recruit T1D case trios and singletons from Pittsburgh.*

In Milestone 1A we proposed to increase the number of participants in our GWA study by applying to the WTCCC for access to their GWA dataset. The WTCC dataset contains SNP genotypes collected from 2,000 T1D and 3,000 non-T1D participants. The WTCCC and our group used identical Affymetrix manufactured genotyping arrays to evaluate risk for T1D. Therefore, SNPs evaluated by the WTCCC and Pittsburgh GWA studies are essentially the same. As a result, the dataset collected at Pittsburgh and by the WTCCC can be combined in order to increase the overall size of the study thus replicating the study while simultaneously improving the power of the study to identify T1D susceptibility alleles. Once the application for data sharing is approved by the WTCCC committee that evaluates requests, our analysis of the WTCCC dataset can begin immediately. The application is complete, pending final signatures from the University of Pittsburgh Office of Grants and Research, and we anticipate submission to the WTCCC will occur during the upcoming research quarter. A copy of the application is attached as an Appendix to this quarterly report.

In Milestone 1B, we proposed to continue recruitment of T1D case trios and singletons from the University of Pittsburgh Medical Center (UPMC). Recruitment of new participants is one of the primary focuses of the project and patients will be recruited continuously during the project's lifetime. The new participants recruited during Goal 1 will be used to confirm the genetic signals identified from the cohort described in Table 1 that was used during our previously completed GWA. We can follow the health of these patients at multiple times during the project since our IRB approved protocol allows us to contact formerly recruited human subjects for the purpose of updating their health information and, when appropriate, collecting a second vial of blood. Table 2 summarizes the current collection of T1D participants recruited in Pittsburgh.

Table 2. Characteristics of Singleton and Family Trios Participants Recruited at Pittsburgh.

	Pittsburgh Cohorts		GoKinD Cohort
	Singletons	Family Trios	Family Trios
Demographic Characteristics:			
Number of Participants	46	26	569
European American (%)	100	100	100
History of Diabetes:			
Type 1 Diabetes (%)	100	100	100
Mean Age at T1D Diagnosis (yr)	14(7)	14(7)	11(6)

Values in parentheses indicate standard deviation from the mean. Pittsburgh samples were recruited at the University of Pittsburgh Medical Center. GoKinD participants were recruited by the Genetics of Kidneys in Diabetes Study.

Goal 2. *Determine high priority SNPs to study.* ***Milestone 2A.*** *Identify confirmed and/or suggestive T1D risk loci using our original GWA study.* ***Milestone 2B.*** *Identify novel loci to evaluate for T1D risk.*

We have completed Milestones 2A by evaluating our T1D GWA data against a list of known and suspected T1D risk loci. The identities of the T1D risk loci were obtained from T1Dbase (http://t1dbase.org) a world-wide-web accessible database of T1D genetic risk information. The genomic location of these confirmed T1D susceptibility loci were compared with the location of SNPs that during the previously completed GWA associated with p-values exceeding 0.001. The analysis identified 12 strongly associated SNPs occurring within 50Kb of T1D candidate loci. SNPs occurring near confirmed T1D loci are summarized in the Table 3.

Table 3. Summary of GWA Results for SNPs located within 50Kb of confirmed T1D Susceptibility Loci.

Chr	dbSNP_ID	Location	Locus	Official Full Name
1	Rs6679677	114,105,331	PHTF1	Putative homeodomain transcription factor 1
2	Rs231726	204,449,111	CTLA4	Cytotoxic T-lymphocyte-associated protein 4
2	Rs1427676	204,449,411	CTLA4	Cytotoxic T-lymphocyte-associated protein 4
7	Rs10232298	127,006,653	PAX4	Paired box 4

10	Rs10905668	6,132,061	IL2RA	Interleukin 2 receptor, alpha
10	Rs10905669	6,132,099	IL2RA	Interleukin 2 receptor, alpha
10	Rs7074372	6,228,597	IL2RA	Interleukin 2 receptor, alpha
12	Rs11171739	54,756,892	ERBB3	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avain)
13	Rs3812863	27,443,268	PDX1	Pancreatic and duodenal homeobox 1
16	Rs9746695	11,115,395	CLEC16A	C-type lectin domain family 16, member A
16	Rs11640295	11,298,229	CLEC16A	C-type lectin domain family 16, member A
20	Rs2425756	44,223,669	CD40	CD40 molecule, TNF receptor superfamily member 5

Signals from the HLA region on chromosome 6 are excluded from the information presented in the table.

Milestone 2B has also been completed during the last quarterly research period. The GWA scan was evaluated to identify SNPs that implicated novel genetic loci for further investigation. As summarized in Table 4 there were 24 loci identified by SNPs that had p-values exceeding the 10th percentile of all association signals garnered during the previously completed GWA. For several loci, such as, KCNH7, STXBP5, and RXFP2 there were clusters of 5 or more SNPs covering genomic regions from 64Kb and up to 247Kb. These loci, as well as others with strong evidence for association are included in Table 4.

Table 4. Summary of GWA Results for SNPs located within 50Kb of Novel Loci.

Chr	dbSNP ID	Location	Locus	Official Full Name
1	rs10858046	114,875,303	AMPD1	Adenosine monophosphate deaminase 1 (isoform M)
1	rs6689415	114,908,992	AMPD1	Adenosine monophosphate deaminase 1 (isoform M)
1	rs2268697	115,029,016	AMPD1	Adenosine monophosphate deaminase 1 (isoform M)
1	rs6427093	165,823,621	CREG1	Cellular repressor of E1A-stimulated genes 1
1	rs2094028	194,417,119	KCNT2	Potassium channel, subfamily T, member 2
1	rs16839682	194,497,622	KCNT2	Potassium channel, subfamily T, member 2
2	rs12470503	163,405,297	KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7
2	rs1001911	163,421,846	KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7
2	rs1492990	163,422,123	KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7
2	rs10203602	163,442,376	KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7
2	rs1364562	163,449,972	KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7
2	rs307869	163,469,837	KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7
2	rs2028201	201,011,255	KCTD18	Potassium channel tetramerisation domain containing 18
2	rs821804	238,928,000	TRAF3IP1	TNF receptor-associated factor 3 interacting protein 1
2	rs523437	239,063,548	TRAF3IP1	TNF receptor-associated factor 3 interacting protein 1
2	rs580209	239,065,169	TRAF3IP1	TNF receptor-associated factor 3 interacting protein 1
5	rs10042443	55,388,527	IL6ST	Interleukin 6 signal transducer (gp130, oncostatin M receptor)
6	rs950286	374,457	IRF4	Interferon regulatory factor 4
6	rs4959880	387,206	IRF4	Interferon regulatory factor 4
6	rs13214605	394,483	IRF4	Interferon regulatory factor 4
6	rs2498587	118,133,432	GOPC	Golgi associated PDZ and coiled-coil motif containing
6	rs9322085	147,354,835	STXBP5	Syntaxin binding protein 5 (tomosyn)
6	rs7765767	147,366,631	STXBP5	Syntaxin binding protein 5 (tomosyn)
6	rs1529017	147,392,995	STXBP5	Syntaxin binding protein 5 (tomosyn)
6	rs6919814	147,412,758	STXBP5	Syntaxin binding protein 5 (tomosyn)
6	rs6920096	147,412,791	STXBP5	Syntaxin binding protein 5 (tomosyn)
6	rs1881660	147,420,052	STXBP5	Syntaxin binding protein 5 (tomosyn)
6	rs6930740	147,438,438	STXBP5	Syntaxin binding protein 5 (tomosyn)
6	rs6920901	147,470,991	STXBP5	Syntaxin binding protein 5 (tomosyn)

7	rs1294610	7,332,182	C1GALT1	Core 1 synthase, glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase, 1
7	rs1294651	7,381,742	C1GALT1	Core 1 synthase, glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase, 1
7	rs1404829	144,255,618	TPK1	Thiamin pyrophosphokinase 1
9	rs2808784	115,961,345	COL27A1	Collagen, type XXVII, alpha 1
9	rs2567724	115,961,754	COL27A1	Collagen, type XXVII, alpha 1
9	rs10982093	115,976,726	COL27A1	Collagen, type XXVII, alpha 1
10	rs2803590	88,362,973	BMPR1A	Bone morphogenetic protein receptor, type 1A
10	rs2803568	88,426,595	BMPR1A	Bone morphogenetic protein receptor, type 1A
11	rs574434	100,384,527	PGR	Progesterone receptor
12	rs7305099	845,537	WNK1	WNK lysine deficient protein kinase 1
12	rs12828016	868,626	WNK1	WNK lysine deficient protein kinase 1
13	rs7328020	30,310,128	ALOX5AP	Arachidonate 5-lipoxygenase-activating protein
13	rs4941903	31,140,471	RXFP2	Relaxin/insulin-like family peptide receptor 2
13	rs7336382	31,146,759	RXFP2	Relaxin/insulin-like family peptide receptor 2
13	rs3950017	31,151,027	RXFP2	Relaxin/insulin-like family peptide receptor 2
13	rs641696	31,376,444	RXFP2	Relaxin/insulin-like family peptide receptor 2
13	rs203415	31,387,796	RXFP2	Relaxin/insulin-like family peptide receptor 2
13	rs1041637	109,307,346	IRS2	Insulin receptor substrate 2
14	rs1873274	62,287,052	KCNH5	Potassium voltage-gated channel, subfamily H (eag-related), member 5
15	rs4779798	27,197,881	APBA2	Amyloid beta (A4) precursor protein-binding, family A, member 2 (X11-like)
15	rs2676083	31,397,903	RYR3	Ryanodine receptor 3
15	rs2596181	31,438,457	RYR3	Ryanodine receptor 3
18	rs9959583	20,993,778	ZNF521	Zinc finger protein 521
18	rs8085678	20,998,828	ZNF521	Zinc finger protein 521
18	rs1971613	21,589,323	ZNF521	Zinc finger protein 521
19	rs687136	45,366,831	AKT2	v-akt murine thymoma viral oncogene homolog 2
19	rs687155	45,366,846	AKT2	v-akt murine thymoma viral oncogene homolog 2

Signals from the HLA region on chromosome 6 are excluded from the information presented in the table.

Goal 3. *Initiate genotyping of priority SNPs.* ***Milestone 3A.*** *Use the cohort of T1D family trios to evaluate confirmed SNPs for T1D risk.* ***Milestone 3B.*** *Use the cohort of T1D family trios to evaluate novel SNPs for T1D risk.*

As outlined in Goal 3 Milestones 3A and 3B we have begun the process of evaluating SNPs selected from those shown in Tables 3 and 4 in order to evaluate interesting SNPs near confirmed T1D loci (Milestone 3A) and SNPs occurring near interesting genes (Milestone 3B) for risk of T1D. The 17 SNPs listed in Table 5 indicate those that have been selected for this step. At this stage of the process we have acquired the necessary reagents for PCR amplification and iPlex (Sequenom, Inc. San Diego, CA) style genotyping of DNA samples. The DNA samples are taken from our repository of affected (T1D) family trios. The data will be evaluated initially by TDT analysis.

Table 5. SNPs selected for iPlex Based Genotyping.

Chr	dbSNP_ID	Location	Locus	Official Full Name
1	rs10858046	114,875,303	AMPD1	Adenosine monophosphate deaminase 1 (isoform M)
1	rs10917913	162,343,920	MGST3	Microsomal glutathione S-transferase 3
1	rs7555369	162,344,746	MGST3	Microsomal glutathione S-transferase 3
1	rs16839682	194,497,622	KCNT2	Potassium channel, subfamily T, member 2
2	rs12470503	163,405,297	KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7

2	rs1001911	163,421,846	KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7
2	rs1492990	163,422,123	KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7
2	rs10203602	163,442,376	KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7
2	rs1364562	163,449,972	KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7
5	rs10042443	55,388,527	IL6ST	Interleukin 6 signal transducer (gp130, oncostatin M receptor)
6	rs13214605	394,483	IRF4	Interferon regulatory factor 4
6	rs7765767	147,366,631	STXBP5	Syntaxin binding protein 5 (tomosyn)
9	rs7864499	91,744,788	IL6RL1	Interleukin 6 receptor-like 1
9	rs10982093	115,976,726	COL27A1	Collagen, type XXVII, alpha 1
12	rs12828016	868,626	WNK1	WNK lysine deficient protein kinase 1
18	rs9959583	20,993,778	ZNF521	Zinc finger protein 521
18	rs8085678	20,998,828	ZNF521	Zinc finger protein 521

12. Statement of Plans and Milestones for the Next Quarter

There are 3 goals that will be the focus of our work during the next research quarter.

Goal 1. Continue to recruit new T1D affected singletons and family trios. **Milestone 1A.** File application for access to the Wellcome Trust Case Control Consortium for access to the T1D (N=2,000) and control (N=3,000) cohorts. **Milestone 1B.** Continue to recruit T1D case trios and singletons from Pittsburgh.

Critical to the success of the project to identify novel loci associated with T1D is the inclusion of as many T1D affected family trios and singletons as possible. To accomplish this goal we will, throughout the lifespan of the project, continue to identify genetic datasets that are available for incorporation into our research (Milestone 1A) and recruit new participants from resources available through our association with UPMC (Milestone 1B). We anticipate that these milestones, or very similar milestones, will be the first goal of every research quarterly report.

Goal 2. Genotype select SNPs using T1D affected family trios. **Milestone 2A.** Using iPlex genotyping methodology genotype SNPs (N=12) occurring near confirmed T1D susceptibility loci (listed in Table 3) in T1D affected family trios (N=595) to measure association at truly positive SNPs. **Milestone 2B.** Beginning with the SNPs (N=17) listed in Table 5 of this quarterly report, genotype available T1D (N=595) affected family trios using iPlex methodology (Sequenom, Inc., San Diego, CA).

We will accomplish Goal 2 by formatting DNA samples from 595 T1D affected family trios (comprised of DNA from father, mother, and affected offspring) into 384-well trays. Sequenom iPlex methodology will be used to genotype the SNPs listed in Tables 3 and 5 for confirmed and novel T1D susceptibility loci, respectively. The iPlex methodology is a mass spectrometer based approach to DNA genotyping and the method is available at the University of Pittsburgh Genomics Core Laboratory. In the event that select SNPs need to be confirmed by an alternative strategy, we will employ TaqMan genotyping methodology that is available in the PI's laboratory.

Goal 3. Analyze the genotyping results garnered during Goal 2 in order to determine which SNPs replicate signal for association with T1D when analyzed in affected family trios. **Milestone 3A.** Evaluate genotyping data from Goal 2 for quality based upon SNP missingness occurring less than 5% of the time, sample missingness occurring less than 5% of the time, and consistency with known family pedigrees. **Milestone 3B.** Perform transmission equilibrium test (TDT) analysis of high quality genotyping data garnered in Goal 2 in order to identify SNPs that replicated in the GWA and TDT analyses.

The Goal 3 will be accomplished using the experimental data garnered during completion of Goal 2. Data will be initially characterized for quality by determined which, if any, samples or SNP genotyping assays result in failure frequency exceeding 5% of the total number of assays performed. Quality control will also include a test for family pedigree consistent with the pattern of Mendelian inheritance between parents and offspring. Datasets that are consistent with parental genotypes and quality control threshold for missingness will be used to determine by TDT analysis p-values associated with T1D susceptibility.

In our third quarterly scientific progress report (03/01/08 – 05/31/08) we then reported the following findings:

Type 1 Diabetes Genome-Wide Association Study Update:
As outlined in the following sections, the research goals for the second quarter of 2008 were 3-fold: 1) continue to recruit new T1D affected singletons and family trios; 2) genotype select SNPs using T1D affected family trios; and 3) analyze the genotyping results in order to determined which SNPs replicate signal for association with T1D. Our progress for each goal is addressed below.

Goal 1. Continue to recruit new T1D affected singletons and family trios. **Milestone 1A.** File application for access to the Wellcome Trust Case Control Consortium for access to the T1D (N=2,000) and control (N=3,000) cohorts. **Milestone 1B.** Continue to recruit T1D case trios and singletons from Pittsburgh.

The overall goal of the project is to identify novel genetic loci that impact risk of developing T1D. Critical to the success of the research is the inclusion of as many T1D affected singletons and family trios as possible. To accomplish this goal we will throughout the lifespan of the project identify publicly available genetic datasets (Milestone 1A) and recruit new participants (Milestone 1B) that can be incorporated into our study. Milestone 1A has recently been completed. On June 9, 2008 we submitted our application to the Wellcome Trust Case Control Consortium (WTCCCC) that, if approved, will allow access to the WTCCC cohort of 2,000 T1D and 3,000 non-T1D singletons. This collection of case and control samples has been genotyped previously using the same methodology that was used during the genome-wide association scan (GWAS) performed in our laboratory. An important outcome of this is that it allows the preexisting WTCCC and Pittsburgh datasets to be combined in order to provide a powerful resource to test association statistics between genetic loci and T1D risk. We anticipate that during the next research quarter we will be notified as to whether our application to the WTCCC has been approved. If this is the case we will immediately begin to compare genetic association signals observed in our data with that observed by the WTCCC.

The second milestone of Goal 1 was to continue to recruit T1D case trios and singletons from Pittsburgh. The updated table for recruited T1D singletons and family trios is illustrated below (Table 1). In addition to participants recruited at Pittsburgh, we have arranged for DNA samples from participants recruited at Ulm, Germany (N=440). The German cohort consists of T1D (N=190) and non-T1D (N=250) singletons. The non-T1D samples add significantly to our study in that they represent a new collection of control samples to which the genetics of T1D case samples can be compared. The N=2,238 samples listed in Table 1 are available as DNA and are now contained within the Pittsburgh repository.

Table 1. T1D Singleton (N=1,988), non-T1D Singletons (N=250), and Family Trio (N=595) Characteristics.

	Pittsburgh Cohort		GoKinD Cohort		German Cohort	
	Singleton (T1D)	Trio (T1D)	Singleton (T1D)	Trio (T1D)	Singleton (T1D)	Singleton (non-T1D)
Demographic Characteristics:						
Number of Participants	55	26	1743	569	190	250
Caucasian (%)	100	100	100	100	100	100
Type 1 Diabetes (%)	100	100	100	100	100	0

Pittsburgh samples were recruited at the University of Pittsburgh Medical Center. GoKinD participants were recruited by the Genetics of Kidneys in Diabetes Study. The German cohort was recruited from Ulm, Germany as part of the EMIL cohort study.

Goal 2. Genotype select SNPs using T1D affected family trios. **Milestone 2A.** Using iPlex genotyping methodology genotype SNPs occurring near confirmed T1D susceptibility loci in T1D affected family trios (N=569) in order to measure association at truly positive SNPs. **Milestone 2B.** Beginning with the novel and clustered SNPs listed in Table 2 of this quarterly report, genotype available T1D (N=569) affected family trios using iPlex methodology.

Genetic loci were identified using results obtained from the T1D GWAS and are listed in Table 2. Loci have been sorted by whether they represent unique signals for T1D risk (i.e., Cluster of Greater Than 6 SNPs and Novel T1D Candidates) or have been identified previously by genetic analysis of preexisting T1D cohorts (i.e., T1DBase Candidate). Among the loci recognized as unique signals were genes encoding the insulin signaling proteins AKT2 and IRS2. Although no previously published genetic screens have associated genetic polymorphisms of these loci with T1D risk, gene knockout and knockdown studies performed in mice have indicated that insulin signaling through IRS2 is critical for pancreatic beta cell survival, and activity of the IRS2/AKT2 dependent insulin signaling pathway has been suggested to play a role in T1D onset. During the next research period, the loci listed in Table 2 will continue to form the basis for choosing which SNPs to test by iPlex genotyping technology.

Goal 3. Analyze genotyping results garnered during Goal 2 in order to determine which SNPs replicate signal for association with T1D when analyzed in affected family trios. **Milestone 3A.** Evaluate genotyping data from Goal 2 for quality based upon SNP missingness occurring less than 5% of the time, sample missingness occurring less than 5% of the time, and consistency with known family pedigrees. **Milestone 3B.** Perform transmission equilibrium test (TDT) analysis of high quality genotyping data garnered in Goal 2 in order to identify SNPs that replicated in the GWA and TDT analyses.

In order to complete the milestones of Goal 3 we have used iPlex genotyping methodology to evaluate a subset of the loci identified during analysis of the T1D GWAS (Table 3). The iPlex methodology is a mass spectrometer based approach to DNA genotyping. The method is available at the University of Pittsburgh Genomics Core Laboratory. The iPlex generated genotyping results have been obtained for 10 loci (Table 3) on DNA samples from 569 T1D affected family trios. As our work into elucidating the genetics of T1D susceptibility progresses we will continue to evaluate genetic loci for association with the T1D phenotype.

Using the criteria stated in Milestones 3A and 3B of Goal 3 we evaluated potential risk loci for genetic association with T1D. The first set of genotyping data has been characterized for quality of the genotyping results. Quality control analyses showed that the first 15 SNPs tested exhibited greater than 95% success for providing genotype data and that greater than 95% of the DNA samples provided genotyping data for each SNP (data not shown). Quality control testing also included the requirement that family pedigrees were consistent with the pattern of Mendelian inheritance of genetic markers between parents and offspring. It was observed that all T1D family trios passed the pedigree test. High quality genotyping results were also examined for accuracy and reproducibility by comparing genotypes determined using TaqMan and iPlex (Table 4). Quality control analysis of the iPlex data indicated that genotyping results were highly reproducible when compared with the identical samples genotyped using TaqMan assays. Taken together, these analyses indicate that SNP typing assays, DNA samples, and the iPlex methodology were of sufficient quality to accomplish the goal of association testing for all interesting genetic markers.

Summarized in Table 5 are the results for genetic association testing of 15 SNPs with the T1D phenotype. The assay used T1D affected family trios (N=569) and was evaluated using the Transmission Disequilibrium Test (TDT). The results indicate that this first set of SNPs did not replicate the results observed during the T1D GWAS of case and control singletons. A likely cause of this discrepancy is that the GWAS data was confounded by population stratification of T1D (case) and non-T1D (control) participants, a common occurrence in case and control analyses. The TDT test is notable for controlling errors that correlate with population stratification. In fact, the overall study design for our research project anticipated this possibility when we chose to follow up on the results of the case and control T1D GWAS with replication using T1D affected family trios and TDT analysis. TDT analysis of additional loci (Table 6) will occur during the next

research period and will confirm whether these other signals, observed as significant during the T1D GWAS, were due to population stratification or reflect true positive signals for genetic association.

Table 2. High value loci identified by genome-wide assocation scan for T1D.

Locus	Classification	Locus	Classification
ADNP2	Cluster of > 6 SNPs	CERKL	T1DBase Candidate
CFDP1	Cluster of > 6 SNPs	CLEC16A	T1DBase Candidate
KHDRBS2	Cluster of > 6 SNPs	CNTN4	T1DBase Candidate
LOC646112	Cluster of > 6 SNPs	CTLA4	T1DBase Candidate
SEMA6D	Cluster of > 6 SNPs	CYBA	T1DBase Candidate
TWISTNB	Cluster of > 6 SNPs	DBP	T1DBase Candidate
AKT2	Novel T1D Candidate	DEPDC2	T1DBase Candidate
ALOX5AP	Novel T1D Candidate	DIAPH3	T1DBase Candidate
AMPD1	Novel T1D Candidate	ELF1	T1DBase Candidate
APBA2	Novel T1D Candidate	ERBB3	T1DBase Candidate
BMPR1A	Novel T1D Candidate	FADD	T1DBase Candidate
C1GALT1	Novel T1D Candidate	FBXL4	T1DBase Candidate
CCK	Novel T1D Candidate	IL2RA	T1DBase Candidate
COL27A1	Novel T1D Candidate	IL6	T1DBase Candidate
CREG1	Novel T1D Candidate	IRF8	T1DBase Candidate
GOPC	Novel T1D Candidate	ITPR3	T1DBase Candidate
IL6RL1	Novel T1D Candidate	LOC728091	T1DBase Candidate
IL6ST	Novel T1D Candidate	LPL	T1DBase Candidate
IRF4	Novel T1D Candidate	MAPK14	T1DBase Candidate
IRS2	Novel T1D Candidate	MED12L	T1DBase Candidate
KCNH5	Novel T1D Candidate	OPCML	T1DBase Candidate
KCNT2	Novel T1D Candidate	PALLD	T1DBase Candidate
KCTD18	Novel T1D Candidate	PAX4	T1DBase Candidate
MGST3	Novel T1D Candidate	PDX1	T1DBase Candidate
PGR	Novel T1D Candidate	PHTF1	T1DBase Candidate
RXFP2	Novel T1D Candidate	PKD1L1	T1DBase Candidate
RYR3	Novel T1D Candidate	POLR2B	T1DBase Candidate
SLC30A8	Novel T1D Candidate	SENP7	T1DBase Candidate
STXBP5	Novel T1D Candidate	SEZ6L	T1DBase Candidate
TPK1	Novel T1D Candidate	SLC7A1	T1DBase Candidate
TRAF3IP1	Novel T1D Candidate	SLCO3A1	T1DBase Candidate
WNK1	Novel T1D Candidate	SOS2	T1DBase Candidate
ZNF521	Novel T1D Candidate	SPG7	T1DBase Candidate
AIRE	T1DBase Candidate	STAT4	T1DBase Candidate
ANKRD15	T1DBase Candidate	TLR2	T1DBase Candidate
AR	T1DBase Candidate	TLR7	T1DBase Candidate
BARD1	T1DBase Candidate	TMEM23	T1DBase Candidate
C4orf31	T1DBase Candidate	TNFSF11	T1DBase Candidate
C9orf25	T1DBase Candidate	TRAFD1	T1DBase Candidate
CAT	T1DBase Candidate	ULBP3	T1DBase Candidate
CD40	T1DBase Candidate	VDR	T1DBase Candidate
CD48	T1DBase Candidate		

Table 3. Novel T1D genetic loci tested using iPlex genotyping.

Locus	Description
AMPD1	Adenosine monophosphate deaminase 1 (isoform M)
COL27A1	Collagen, type XXVII, alpha 1
IL6RL1	Interleukin 6 receptor-like 1
IL6ST	Interleukin 6 signal transducer (gp130, oncostatin M receptor)
IRF4	Interferon regulatory factor 4
KCNH7	Potassium voltage-gated channel, subfamily H (eag-related), member 7
MGST3	Microsomal glutathione S-transferase 3
STXBP5	Syntaxin binding protein 5 (tomosyn)
WNK1	WNK lysine deficient protein kinase 1
ZNF521	Zinc finger protein 521

Table 4. Quality control analysis for iPlex and TaqMan genotyping of rs8078936 and rs222747

Genotype Based Typing:	iPlex Genotyping			TaqMan Genotyping			Chi Square
dbSNP ID	AA	AB	BB	AA	AB	BB	p-value
rs8078936	394	1423	1252	391	1433	1252	0.88
rs222747	182	1170	1735	179	1151	1732	0.78

Allele Based Typing:	iPlex Genotyping		TaqMan Genotyping		
dbSNP ID	Allele A	Allele B	Allele A	Allele B	p-value
rs8078936	2211	3927	2215	3937	0.98
rs222747	1534	4640	1509	4615	0.79

p-values were calculated using chi square analysis and 1 degree of freedom.

Table 5. TDT results from T1D affected family trios (N=569)

				Transmission Disequilibrium Test Data			
dbSNP ID	Chr	Location	Locus	Allele 1	Allele 2	Total	p-value
rs10858046	1	114,875,303	AMPD1	239	235	474	0.85
rs10917913	1	162,343,920	MGST3	283	264	547	0.42
rs7555369	1	162,344,746	MGST3	263	281	544	0.44
rs12470503	2	163,405,297	KCNH7	236	257	493	0.34
rs1001911	2	163,421,846	KCNH7	255	225	480	0.17
rs1492990	2	163,422,123	KCNH7	254	226	480	0.20
rs10203602	2	163,442,376	KCNH7	197	204	401	0.73
rs1364562	2	163,449,972	KCNH7	199	205	404	0.77
rs10042443	5	55,388,527	IL6ST	226	238	464	0.58
rs13214605	6	394,483	IRF4	175	183	358	0.67
rs7765767	6	147,366,631	STXBP5	248	267	515	0.40
rs7864499	9	91,744,788	IL6RL1	225	213	438	0.57
rs10982093	9	115,976,726	COL27A1	112	112	224	1.00
rs12828016	12	868,626	WNK1	238	248	486	0.65
rs8085678	18	20,998,828	ZNF521	272	265	537	0.76

Table 6. High priority SNPs for TDT analysis during the upcoming research quarter.

dbSNP ID	Chr	Location	Locus	dbSNP ID	Chr	Location	Locus
rs6689415	1	114,908,992	AMPD1	rs10825557	10	51,709,687	TMEM23
rs2268697	1	115,029,016	AMPD1	rs2803590	10	88,362,973	BMPR1A
rs6427093	1	165,823,621	CREG1	rs11032755	11	34,533,161	CAT
rs2094028	1	194,417,119	KCNT2	rs574434	11	100,384,527	PGR
rs1441147	2	182,220,101	CERKL	rs7305099	12	845,537	WNK1
rs2696344	2	182,241,456	CERKL	rs7484827	12	46,409,830	VDR
rs2028201	2	201,011,255	KCTD18	rs1170188	13	41,576,971	TNFSF11
rs523437	2	239,063,548	TRAF3IP1	rs1170187	13	41,577,835	TNFSF11
rs580209	2	239,065,169	TRAF3IP1	rs670676	13	41,599,739	TNFSF11
rs9877118	3	2,578,450	CNTN4	rs1041637	13	109,307,346	IRS2
rs6771904	3	152,630,421	MED12L	rs7155762	14	49,487,928	SOS2
rs7654918	4	58,045,356	POLR2B	rs7155791	14	49,487,980	SOS2
rs343173	4	121,980,468	C4orf31	rs2676083	15	31,397,903	RYR3
rs950286	6	374,457	IRF4	rs16959754	15	45,680,428	SEMA6D
rs2498587	6	118,133,432	GOPC	rs10781976	16	73,872,339	CFDP1
rs6920096	6	147,412,791	STXBP5	rs1559339	16	73,876,147	CFDP1
rs6920901	6	147,470,991	STXBP5	rs4888405	16	73,985,697	CFDP1
rs6922684	6	150,491,727	ULBP3	rs11076678	16	87,165,397	CYBA
rs4870524	6	150,820,890	ULBP3	rs7184960	16	88,489,162	SPG7
rs1294610	7	7,332,182	C1GALT1	rs1971613	18	21,589,323	ZNF521
rs1404829	7	144,255,618	TPK1	rs1064059	18	75,994,674	ADNP2
rs2380437	8	69,156,370	DEPDC2	rs3786803	19	35,655,444	LOC728091
rs7032335	9	34,572,883	C9orf25	rs687136	19	45,366,831	AKT2
rs2808784	9	115,961,345	COL27A1	rs687155	19	45,366,846	AKT2
rs2567724	9	115,961,754	COL27A1	rs281408	19	53,925,218	DBP
				rs12975781	19	53,941,510	DBP

12. Statement of Plans and Milestones for the Next Quarter

There are 3 goals that will be the focus of our work during the next research quarter.

Goal 1. Continue to recruit new T1D affected singletons and family trios. **Milestone 1A.** As a continuing goal of the project into the genetics of T1D we will pursue the recruitment of new T1D case trios and singletons from Pittsburgh. **Milestone 1B.** Continue to arrange new collaborations with researchers at other institutions (e.g., Wellcome Trust Case Control Consortium and EMIL researchers at Ulm, Germany) to gain increased access to DNA repositories and datasets useful for evaluating the genetics of T1D

Goal 2. Continue to genotype select SNPs using T1D affected family trios. **Milestone 2A.** Select loci from those listed in Table 6 for use during genotyping replication studies. **Milestone 2B.** Using the T1D affected family trios determine the genotype of each SNP chosen during Milestone 2A.

Goal 3. Analyze the genotyping results garnered during Goal 2 in order to determine which SNPs replicate signal for association with T1D when analyzed using affected family trios. **Milestone 3A.** Analyze the results from Goal 2 by Transmission Disequilibrium Testing to identify loci that reproduce the results observed during the T1D GWAS study.

In the fourth and final quarterly scientific progress report (06/01/08 - 08/26/08) of year 01, we now report on our cumulative results.

Type 1 Diabetes Genome-Wide Association Study Update:
As outlined in the following sections, the research goals for the third quarter of 2008 were 3-fold: 1) continue to recruit new Type 1 Diabetes (T1D) affected singletons and family trios; 2) continue to genotype select SNPs using T1D affected family trios; and 3) analyze the genotyping results garnered during Goal 2 in order to determine which SNPs replicate signal for association with T1D when analyzed using affected family trios. Our progress for each goal is addressed below.

Goal 1. Continue to recruit new T1D affected singletons and family trios. **Milestone 1A.** As a continuing goal of the project into the genetics of T1D we will pursue the recruitment of new T1D case trios and singletons from Pittsburgh. **Milestone 1B.** Continue to arrange new collaborations with researchers at other institutions (e.g., Wellcome Trust Case Control Consortium and EMIL researchers at Ulm, Germany) to gain increased access to DNA repositories and datasets useful for evaluating the genetics of T1D.

The first and second milestones of Goal 1 were to continue to recruit T1D case trios and singletons from Pittsburgh and to actively pursue collaborations with other research centers to increase the samples available to our study. The updated table for recruited T1D singletons and family trios is illustrated below indicating that during the last research quarter an additional 3 participants were recruited in Pittsburgh (Table 1). In contrast, participants recruited at Ulm, Germany have increased substantially from N=440 reported previously to N=2,608 during the last research quarter. The German cohort now consists of T1D (N=858) and non-T1D (N=1,750) singletons. The non-T1D samples add significantly to our study in that they represent a collection of control samples to which the genetics of T1D case samples can be compared. As of the end of the present research quarter, the N=4,978 samples listed in Table 1 are available as DNA and are now contained within the Pittsburgh repository.

Milestones 1A and 1B of Goal 1 represent goals that will be pursued continuously throughout the project. For the recently completed research quarter the milestones were accomplished upon recruitment of additional participants from Pittsburgh and, more significantly, upon receipt of greater than 2,000 DNA samples from our collaborator in Germany.

Table 1. T1D Singleton (N=2,659), non-T1D Singletons (N=1,750), and T1D Family Trio (N=595) Characteristics.

	Pittsburgh Cohort		GoKinD Cohort		German Cohort	
	Singleton (T1D)	Trio (T1D)	Singleton (T1D)	Trio (T1D)	Singleton (T1D)	Singleton (non-T1D)
Demographic Characteristics:						
Number of Participants	58	26	1,743	569	858	1,750
Caucasian (%)	100	100	100	100	100	100
Type 1 Diabetes (%)	100	100	100	100	100	0

Pittsburgh samples were recruited at the University of Pittsburgh Medical Center. GoKinD participants were recruited by the Genetics of Kidneys in Diabetes Study. The German cohort was recruited from Ulm, Germany as part of the EMIL cohort study.

Goal 2. Continue to genotype select SNPs using T1D affected family trios. **Milestone 2A.** Select loci from those listed in Table 6 (of the previous quarterly report) for use during genotyping replication studies. **Milestone 2B.** Using the T1D affected family trios determine the genotype of each SNP chosen during Milestone 2A.

As indicated in Table 2, select SNPs were chosen for analysis using mass spectrometry based genotyping method developed by Sequenom, Inc (San Diego, CA). SNPs were chosen based upon the signal observed during the initial genome-wide association (GWA) scan for T1D that has been described in preceding quarterly reports. In total, 21 SNPs were genotyped using the Sequenom iPlex method and high quality data was returned on 19, providing an overall return of high quality data from SNPs sent for analysis of 90%. The 2 SNPs that failed to provide genotyping data were rs16839682 (*KCNT2*) and rs9959583 (*ZNF521*) and were due to lack of signal indicating that these samples may have failed to amplify during sample preparation. Subsequent analysis of the genotyping data obtained from the 19 successfully genotyped markers indicated that family pedigree remained intact and that missing data due to failure to genotype remained below 5% (data not shown).

Successful completion of Milestones 2A and 2B for the current research quarter was achieved upon creation of the genotyping data for these SNPs (Milestone 2A) and quality control analysis for missingness and family pedigree (Milestone 2B).

Table 2. List of SNPs for analysis of association with T1D.

dbSNP ID	Chr	Location	Candidate Locus	Cytogenetic Band	Distance (kb)[1]
rs10858046	1	114,875,303	*AMPD1*	1p13	141.9
rs10917913	1	162,343,920	*MGST3*	1q23	1,523.2
rs7555369	1	162,344,746	*MGST3*	1q23	1,522.3
rs12470503	2	163,405,397	*KCNH7*	2q24.2	1.9
rs1001911	2	163,421,846	*KCNH7*	2q24.2	18.4
rs1492990	2	163,422,123	*KCNH7*	2q24.2	18.6
rs10203602	2	163,442,376	*KCNH7*	2q24.2	38.9
rs1364562	2	163,449,972	*KCNH7*	2q24.2	46.5
rs10042443	5	55,388,527	*IL6ST*	5q11	62.0
rs13214605	6	394,483	*IRF4*	6p25-p23	38.3
rs7765767	6	147,366,631	*STXBP5*	6q24.3	199.9
rs7864499	9	91,744,788	*IL6RL1*	9q22.2	138.5
rs10982093	9	115,976,726	*COL27A1*	9q32	intron
rs12828016	12	868,626	*WNK1*	12p13.3	nonsynonymous
rs877610	17	3,422,240	*TRPV1*	17p13.3	nonsynonymous
rs8078936	17	3,427,145	*TRPV1*	17p13.3	intron
rs161393	17	3,436,263	*TRPV1*	17p13.3	intron
rs222747	17	3,439,949	*TRPV1*	17p13.3	nonsynonymous
rs8085678	18	20,998,828	*ZNF521*	18q11.2	intron

1. Distance between SNP and candidate locus if extragenic. If SNP is intragenic than its funtional position is indicated.

Goal 3. Analyze the genotyping results garnered during Goal 2 in order to determine which SNPs replicate signal for association with T1D when analyzed using affected family trios. **Milestone 3A.** Analyze the results from Goal 2 by Transmission Disequilibrium Testing to identify loci that reproduce the results observed during the T1D GWAS study.

Transmission disequilibrium testing (TDT) of the genotype data obtained during completion of Goal 2 is the focus of the third goal for the current research quarter. The goal was completed by calculating the p-value of association correlated with the 19 SNPs listed in Table 3 in order to estimate the risk of developing T1D. The analysis was based upon genotyping data obtained from 569 family trios in which the offspring was T1D positive. Of the 19 SNPs examined signals for inheritance of one allele over the other, and thus an indication that alleles associate with T1D, varied between 50% (or essentially random) for locus *COL27A1* and 53.7% for SNP rs161393 located within the *TRPV1* locus. Analysis of the genotyping data for rs161393 indicated that out of 486 heterozygous parents there were 225 offspring inheriting allele A and 261 that inherited allele B. This corresponded to a chi square value of 2.7 and p-value of 0.1. However, the criteria of our study has been that for SNPs to be considered as exhibiting significant association with T1D the p-values must be less than 0.05. Thus, for the SNPs examined so far for association with T1D the conclusion is that they are not likely to significantly impact disease risk.

Completion of Goal 3 was accomplished by analysis of the genotyping data obtained during the recently completed research quarter. The results failed to identify novel T1D loci but the analysis did confirm that the genotyping method and laboratory procedures in place in Pittsburgh are sufficient to routinely obtain high quality genetic information from at large number of samples (N=1,707) and on a large number of genetic markers (N=19). The operational infrastructure that has been developed to accomplish the goals of the 2007 and 2008 research period will be used to investigate genetic influence of diabetes complications, while continuing to investigate candidate markers for T1D, in the upcoming research period(s).

Table 3. Transmission disequilibrium analysis of T1D affected family trios.

dbSNP ID	Chr	Location	Locus	Allele 1	Allele 2	Total	%Trans	Chi Sq	p-value
rs10858046	1	114,875,303	*AMPD1*	239	235	474	50.4	0.0	8.5E-01
rs10917913	1	162,343,920	*MGST3*	283	264	547	51.7	0.7	4.2E-01

rs7555369	1	162,344,746	*MGST3*	263	281	544	48.3	0.6	4.4E-01
rs12470503	2	163,405,397	*KCNH7*	236	257	493	47.9	0.9	3.4E-01
rs1001911	2	163,421,846	*KCNH7*	255	225	480	53.1	1.9	1.7E-01
rs1492990	2	163,422,123	*KCNH7*	254	226	480	52.9	1.6	2.0E-01
rs10203602	2	163,442,376	*KCNH7*	197	204	401	49.1	0.1	7.3E-01
rs1364562	2	163,449,972	*KCNH7*	199	205	404	49.3	0.1	7.7E-01
rs10042443	5	55,388,527	*IL6ST*	226	238	464	48.7	0.3	5.8E-01
rs13214605	6	394,483	*IRF4*	175	183	358	48.9	0.2	6.7E-01
rs7765767	6	147,366,631	*STXBP5*	248	267	515	48.2	0.7	4.0E-01
rs7864499	9	91,744,788	*IL6RL1*	225	213	438	51.4	0.3	5.7E-01
rs10982093	9	115,976,726	*COL27A1*	112	112	224	50.0	0.0	1.0E+00
rs12828016	12	868,626	*WNK1*	238	248	486	49.0	0.2	6.5E-01
rs877610	17	3,422,240	*TRPV1*	45	47	92	48.9	0.0	8.3E-01
rs8078936	17	3,427,145	*TRPV1*	227	259	486	46.7	2.1	1.5E-01
rs161393	17	3,436,263	*TRPV1*	225	261	486	46.3	2.7	1.0E-01
rs222747	17	3,439,949	*TRPV1*	197	208	405	48.6	0.3	5.8E-01
rs8085678	18	20,998,828	*ZNF521*	272	265	537	50.7	0.1	7.6E-01

Statement of Plans for the First Quarterly Report of the 2008 to 2009 research period

The goals listed below based upon Task 2 (recruit participants with T1D or T1DN) and Task 3 (genotype candidate genetic markers for association with T1DN) of the overall research project and will use our available cohort of T1DN (cases) and T1D (controls) to perform an additional series of genetic analyses based upon the GWA study that used the 500,000 SNP typing microarray previously reported. The overall goal will be to confirm the strongest genetic signals associated with diabetes complications and will focus upon T1D nephropathy (T1DN).

There are 3 goals that will be the focus of our work during the next research quarter.

Goal 1. Expand genetic analysis to include (T1DN) phenotype. **Milestone 1A.** Analyze available anthropometric data linked with T1D cohort for presence of diabetes complications. **Milestone 1B.** Determine the size of the available cohort for T1D nephropathy (T1DN) and T1D with healthy kidney function.

Goal 2. Garner genetic data to test for association of select SNPs with T1DN. **Milestone 2A.** Use the T1DN and T1D cohort identified during Goal 1 to initiate genotyping of candidate SNPs for their association with T1DN.

Goal 3. Analyze the results of genotyping obtained using the T1DN cohort. **Milestone 3A.** Using T1DN family trios calculate the p-value for association of candidate SNPs with the T1DN phenotype. **Milestone 3B.** Using T1DN and T1D singletons calculate the p-value for association of select SNPs with the T1DN phenotype.

KEY RESEARCH ACCOMPLISHMENTS:

1. Creation of a sample repository containing DNA from 3,170 T1D and 1,750 non-T1D participants

2. Establishment of laboratory methods for multiplex genotyping of 30 or more SNPs simultaneously from as many as 5,000 samples

3. Initiation of genetic studies to evaluate candidate genetic markers for association with diabetes complications

4. Analysis of genetic data to evaluate candidate genes for association with diabetes phenotypes

5. Publication of 9 manuscripts

REPORTABLE OUTCOMES:

Manuscripts (9 publications)

1. Ringquist, S., Nichol, L., Trucco, M. Transplantation Genetics. (2007) *In: Emery and Rimoin's Principles and Practice of Medical Genetics*. eds. D.L. Rimoin, M. Conner, R.E. Pyeritz, B.R. Korf, and A.E. Emery 5th Edition, Elsevier Books, Oxford.

2. Ringquist, S., Pecoraro, C., Trucco, M. (2007) Web-based program for pyrosequencing primer design. ASHI Quarterly 31:50-52.

3. Ringquist S, Pecoraro C, Lu Y, Styche A, Rudert WA, Benos PV, Trucco M. (2007) Web-based primer design software for genome-scale genotyping by pyrosequencing. Methods Mol Biol 373:25-38.

4. Ringquist S, Styche A, Rudert WA, Trucco M. (2007) Pyrosequencing-based strategies for improved allele typing of human leukocyte antigen loci. Methods Mol Biol 373:115-134.

5. Pasquali L, Bedeir A, Ringquist S, Styche A, Bhargava R, Trucco G. (2007) Quantification of CpG island methylation in progressive breast lesions from normal to invasive carcinoma. Cancer Lett 257:136-144.

6. Pasquali L, Trucco M, Ringquist S. (2007) Navigating pathways affecting type 1 diabetic kidney disease. Pediatr Diabetes 8:307-322.

7. Luca D, Ringquist S, Klei L, Lee AB, Gieger C, Wichmann HE, Schreiber S, Krawczak M, Lu Y, Styche A, Devlin B, Roeder K, Trucco M. (2008) On the use of general control samples for genome-wide association studies: genetic matching highlights causal variants. Am J Hum Genet 82:453-463.

8. Zhang L, Perdomo G, Kim DH, Qu S, Ringquist S, Trucco M, Dong HH. (2008) Proteomic analysis of fructose-induced fatty liver in hamsters. Metabolism 57:1115-1124.

9. Lu, L., Boehm, J., Nichol, L., Trucco, M., and Ringquist, S. (2008) Multiplex HLA typing by pyrosequencing. Methods in Molecular Biology (in press).

Abstracts
None

Presentations
None

Patents and Licenses Applied for and/or Issued
None

Degrees Obtained that are Supported by this Award
None

Development of Cell Lines, Tissue or Serum Repositories
1. Repository of DNA samples collected from T1D and T1DN patients exceeding 5,000 samples.

Informatics such as Databases and Animal Models, etc
None

Funding Applied for Based on Work Supported by this Award
None

Employment or Research Opportunities Applied for and/or Received Based on Experience/Training Supported by this Award
None

CONCLUSION:
The principal conclusions for the first year of funding are that: 1) we have been successful in recruiting and obtained via collaboration a large cohort (nearly 5,000 participants) for analysis of diabetes complications and 2) that laboratory methods are in place and have been tested for analysis of large number of samples for association between SNP and T1D related phenotypes. Our work during the first year primarily focused on increasing the cohort size and on validating laboratory methods for genetic analysis. Now that these have been accomplished we will begin to test our principle hypothesis that risk for diabetes complications is dependent upon genes and that environmental factors such as smoking will influence that risk.

The work generated thus far has resulted in 9 publications that are listed under the section entitled "REPORTABLE OUTCOMES". These scientific manuscripts have been published in peer review publications. We are currently editing a new manuscript describing our work on diabetes complications and anticipate having that paper ready for submission to a scientific journal prior to the end of 2008.

The So What Section. What are the implications of this research? Diabetes affects 16 million Americans and 800,000 new cases annually. African, Hispanic, Native and Asian Americans are particularly susceptible to its most severe complications. Costs associated with diabetes may be as high as $132 billion. Diabetes accounts for 42% of new cases of ESRD with over 100,000 cases per year at an average cost of $55,000 per patient annually.

What are the military significance and public purpose of this research? As the military is a reflection of the US population, improved prediction of risk for developing diabetes and diabetic complications among active duty members of the military, their families, and retired military personnel will potentially allow focused preventative treatment of at-risk individuals, providing significant healthcare savings and improved patient well being.

REFERENCES:
1. Ringquist, S., Nichol, L., Trucco, M. Transplantation Genetics. (2007) *In: Emery and Rimoin's Principles and Practice of Medical Genetics*. eds. D.L. Rimoin, M. Conner, R.E. Pyeritz, B.R. Korf, and A.E. Emery 5th Edition, Elsevier Books, Oxford.

2. Ringquist, S., Pecoraro, C., Trucco, M. (2007) Web-based program for pyrosequencing primer design. ASHI Quarterly 31:50-52.

3. Ringquist S, Pecoraro C, Lu Y, Styche A, Rudert WA, Benos PV, Trucco M. (2007) Web-based primer design software for genome-scale genotyping by pyrosequencing. Methods Mol Biol 373:25-38.

4. Ringquist S, Styche A, Rudert WA, Trucco M. (2007) Pyrosequencing-based strategies for improved allele typing of human leukocyte antigen loci. Methods Mol Biol 373:115-134.

5. Pasquali L, Bedeir A, Ringquist S, Styche A, Bhargava R, Trucco G. (2007) Quantification of CpG island methylation in progressive breast lesions from normal to invasive carcinoma. Cancer Lett 257:136-144.

6. Pasquali L, Trucco M, Ringquist S. (2007) Navigating pathways affecting type 1 diabetic kidney disease. Pediatr Diabetes 8:307-322.

7. Luca D, Ringquist S, Klei L, Lee AB, Gieger C, Wichmann HE, Schreiber S, Krawczak M, Lu Y, Styche A, Devlin B, Roeder K, Trucco M. (2008) On the use of general control samples for genome-wide association studies: genetic matching highlights causal variants. Am J Hum Genet 82:453-463.

8. Zhang L, Perdomo G, Kim DH, Qu S, Ringquist S, Trucco M, Dong HH. (2008) Proteomic analysis of fructose-induced fatty liver in hamsters. Metabolism 57:1115-1124.

9. Lu, L., Boehm, J., Nichol, L., Trucco, M., and Ringquist, S. (2008) Multiplex HLA typing by pyrosequencing. Methods in Molecular Biology (in press).

APPENDICES:

We have included the following as appendix materials:

Appendix 1: T1D GWAS p-values

Appendix 2: Wellcome Trust Information

We have also included the following published manuscripts:

Ringquist, S., Pecoraro, C., Trucco, M. (2007) Web-based program for pyrosequencing primer design. ASHI Quarterly 31:50-52.

Ringquist S, Pecoraro C, Lu Y, Styche A, Rudert WA, Benos PV, Trucco M. (2007) Web-based primer design software for genome-scale genotyping by pyrosequencing. Methods Mol Biol 373:25-38.

Pasquali L, Trucco M, Ringquist S. (2007) Navigating pathways affecting type 1 diabetic kidney disease. Pediatr Diabetes 8:307-322.

Luca D, Ringquist S, Klei L, Lee AB, Gieger C, Wichmann HE, Schreiber S, Krawczak M, Lu Y, Styche A, Devlin B, Roeder K, Trucco M. (2008) On the use of general control samples for genome-wide association studies: genetic matching highlights causal variants. Am J Hum Genet 82:453-463.

SUPPORTING DATA:
None

T1D GWAS p-values <0.001 (Revised 122807)

T1D GWAS Data								*T1DBase Candidate Genes*					*Criteria for Candidate SNP*		*Confirmed/Suggested*	
Chr	Location	dbSNP_ID	Dominant	Additive	Recessive	Distance1	QC Summary	Locus	Location1	Location2	Association	Population	Strong p-values	3 SNP within 50kb	T1D Loci	All Loci
1	3,127,816	rs12075867	2.2E-04	1.1E-02	5.4E-04	1,009,026										
1	4,136,842	rs4233262	5.9E-02	3.9E-03	7.1E-04	3,879										
1	4,140,721	rs12074230	4.0E-02	3.4E-03	8.6E-04	551,323										
1	4,692,044	rs3766964	6.0E-04	1.6E-02	1.1E-03	6,214,949										
1	10,906,993	rs1280985	2.2E-04	2.0E-03	1.4E-03	814,822										
1	11,721,815	rs6540992	3.2E-04	2.8E-02	2.1E-01	4,690,349										
1	16,412,164	rs12751725	4.5E-10	6.8E-07	5.4E-08	2,045,443	Poor						ARHGEF19			ARHGEF19
1	18,457,607	rs223185	1.6E-27	0.0E+00	9.8E-15	5,547,285	Fail						IGSF21			IGSF21
1	24,004,892	rs2256179	2.8E-02	4.0E-03	8.4E-04	8,762,281										
1	32,767,173	rs10914574	5.0E-03	1.2E-02	5.4E-04	1,689,877										
1	34,457,050	rs705207	1.9E-02	9.8E-04	1.0E-01	4,107,505										
1	38,564,555	rs913863	8.0E-04	8.4E-04	3.8E-05	59,029	Pass						Intergenic			Intergenic
1	38,623,584	rs7534892	1.7E-02	4.5E-04	1.3E-04	676								676		Intergenic
1	38,624,260	rs6678679	2.0E-02	4.6E-04	1.0E-04	14,443								14,443		Intergenic
1	38,638,703	rs4970590	1.8E-02	6.3E-04	3.4E-03	358,739										Intergenic
1	38,997,442	rs695023	2.2E-02	9.8E-03	6.4E-04	929,779										
1	39,927,221	CG01006038	3.1E-03	1.7E-04	1.3E-02	14,231										
1	39,941,452	rs17561616	3.6E-04	3.4E-03	8.3E-02	4,094,550										
1	44,036,002	rs12401813	7.4E-04	1.0E-01	8.3E-03	10,532,245										
1	54,568,247	rs17110430	3.2E-09	5.9E-05	1.9E-05	1,160,672	Fail						SSBP3			SSBP3
1	55,728,919	rs10736392	3.0E-02	6.7E-04	7.7E-02	28,602								28,602		Intergenic
1	55,757,521	rs12757934	1.2E-02	3.0E-04	2.8E-02	215								215		Intergenic
1	55,757,736	rs4927247	9.2E-03	2.2E-04	2.7E-02	5,284								5,284		Intergenic
1	55,763,020	rs754138	6.0E-03	3.1E-04	2.3E-02	2,578								2,578		Intergenic
1	55,765,598	rs1979827	9.5E-03	4.9E-04	4.0E-02	85								85		Intergenic
1	55,765,683	rs6588561	8.8E-03	4.8E-04	4.0E-02	2,203								2,203		Intergenic
1	55,767,886	rs1995229	1.1E-02	6.2E-04	3.5E-02	1,642,000										Intergenic
1	57,409,886	rs6680219	9.2E-04	2.0E-02	6.9E-03	9,826		DAB1	57,236,167	58,488,799						DAB1
1	57,419,712	rs7548633	8.9E-04	2.2E-02	8.1E-03	1,099,192		DAB1	57,236,167	58,488,799						DAB1
1	58,518,904	rs6668411	1.3E-08	3.3E-07	2.8E-06	1,144,974	Poor	DAB1	57,236,167	58,488,799			DAB1			DAB1
1	59,663,878	rs10493264	7.8E-04	1.1E-02	3.5E-02	1,501,319										
1	61,165,197	rs7530011	9.9E-04	2.7E-03	3.9E-03	2,136,484										
1	63,301,681	rs17124480	9.9E-03	3.8E-02	3.1E-04	1,390,870										
1	64,692,551	rs305543	3.7E-02	3.9E-02	5.9E-04	2,919,610										
1	67,612,161	rs3828069	5.7E-03	9.3E-03	9.8E-04	2,660,339		IL12RB2	67,545,635	67,635,171	Weak	Japanese				IL12RB2
1	70,272,500	rs17131142	7.0E-04	4.6E-04	1.0E-02	11,101,258		LRRC7	69,998,446	70,360,514						LRRC7
1	81,373,758	rs1146414	1.3E-06	2.6E-04	1.4E-03	486,242	Fail						LOC646538			LOC646538
1	81,860,000	rs7523717	1.5E-01	3.6E-03	8.7E-04	6,191,759										
1	88,051,759	rs2117353	2.2E-04	3.9E-04	1.0E-02	2,547,223										
1	90,598,982	rs1156495	2.3E-01	8.7E-02	7.2E-04	23,145										
1	90,622,127	rs4658158	2.3E-01	2.7E-02	5.5E-04	686,946										
1	91,309,073	rs164989	2.1E-04	2.9E-03	1.3E-02	3,795,023										
1	95,104,096	rs12038630	2.4E-03	7.6E-04	3.5E-02	14,056,721										
1	109,160,817	rs1277215	2.4E-02	1.3E-03	5.5E-04	2,566,055										
1	111,726,872	rs4839130	1.1E-01	1.2E-02	4.5E-04	2,157,583										
1	113,884,455	rs10858002	2.2E-03	3.4E-04	6.2E-03	161,245										
1	114,045,700	rs1230649	1.2E-02	7.2E-04	1.0E-01	59,631		PHTF1	114,041,360	114,102,879					PHTF1	PHTF1
1	114,105,331	rs6679677	3.3E-05	1.0E-05	4.4E-02	111,560	Pass	PHTF1	114,041,360	114,102,879			PHTF1		PHTF1	PHTF1
1	114,216,891	rs2488457	3.8E-03	7.1E-04	1.2E-01	658,412		PTPN22	114,157,960	114,215,857	Mixed	Finnish, German, Dutch, Canadian Caucasian, Colombian			PTPN22	PTPN22
1	114,875,303	rs10858046	4.6E-03	7.2E-04	3.0E-01	33,689										
1	114,908,992	rs6689415	6.8E-03	2.2E-04	2.1E-01	120,024										
1	115,029,016	rs2268697	3.4E-03	7.6E-05	8.3E-02	45,079,815	Pass						AMPD1			AMPD1
1	160,108,831	rs10918104	9.1E-04	3.0E-03	1.7E-03	1,085,715										
1	161,194,546	rs10917653	3.1E-02	7.3E-04	3.5E-03	165,460										
1	161,360,006	rs2841989	8.5E-04	9.8E-03	3.9E-03	3,978,593										
1	165,338,599	rs9287077	1.0E-03	8.5E-04	1.5E-02	441,748										
1	165,780,347	rs12042693	1.8E-02	8.5E-04	6.6E-02	474										
1	165,780,821	rs12022039	1.7E-02	8.5E-04	3.2E-02	107,673										
1	165,888,494	rs7527985	5.8E-02	6.0E-03	2.2E-04	16,361,287										
1	182,249,781	rs4651164	6.1E-04	4.6E-03	3.6E-02	8,528,025										
1	190,777,806	rs10489868	1.1E-02	3.2E-04	1.9E-01	3,719,816										
1	194,497,622	rs16839682	6.3E-03	1.1E-04	3.7E-03	5,190,083										
1	199,687,705	rs7553988	2.2E-01	9.7E-04	8.3E-03	1,119								1,119		PHLDA3
1	199,688,824	rs10920205	2.9E-01	3.4E-04	2.3E-02	1,739								1,739		PHLDA3
1	199,690,563	rs3887566	3.8E-01	5.6E-04	2.9E-02	61,045										PHLDA3
1	199,751,608	rs4518935	1.8E-08	2.3E-06	2.6E-05	1,144,760	Fail						PHLDA3			PHLDA3
1	200,896,368	rs7517181	6.1E-04	5.4E-02	6.7E-02	4,299,521										
1	205,195,889	rs10863962	3.2E-01	7.7E-02	9.1E-04	1,036,151										
1	206,232,040	rs684431	6.7E-02	9.7E-04	9.2E-03	2,742,046										
1	208,974,086	rs1890845	8.3E-04	1.2E-02	8.0E-02	7,648,834										
1	216,622,920	rs10482751	8.6E-03	6.2E-04	1.3E-03	1,312										
1	216,624,232	rs4335431	1.7E-03	6.1E-04	1.5E-02	1,521,013										
1	218,145,245	rs2378330	1.3E-03	6.7E-04	8.0E-02	39,228										
1	218,184,473	rs2647442	6.6E-04	1.2E-03	5.8E-02	261,653										
1	218,446,126	rs4142940	9.3E-04	6.1E-03	5.5E-02	4,502,325										
1	222,948,451	rs1436171	8.7E-04	2.7E-03	1.4E-01	6,127								6,127		CNIH3
1	222,954,578	rs1561110	7.1E-04	4.7E-03	1.6E-01	2,565								2,565		CNIH3

1	222,957,143	rs1965776	6.5E-04	3.4E-03	1.9E-01	1,681,099									CNIH3
1	224,638,242	rs2117114	1.5E-01	7.5E-02	5.0E-04	6,273,622		PARP1	224,615,129	224,662,414	Weak	French, French Caucasian			PARP1
1	230,911,864	rs7531764	6.8E-05	2.7E-03	8.6E-04	369,006	Pass					Intergenic			Intergenic
1	231,280,870	rs6684434	6.9E-04	4.1E-03	5.7E-04	5,728,342									
1	237,009,212	rs2689125	1.4E-03	4.6E-03	7.8E-04	1,668,253									
1	238,677,465	rs12410581	4.2E-04	9.9E-04	7.3E-02	392,560									
1	239,070,025	rs12116473	6.5E-04	2.0E-01	1.9E-01	326,992									
1	239,397,017	rs4660036	3.3E-02	1.1E-02	5.5E-04	681,991									
1	240,079,008	rs4149852	2.4E-02	7.4E-04	2.1E-02	3,918,864									
1	243,997,872	rs1316384	7.3E-04	2.0E-02	9.9E-02	34,302									
1	244,032,174	rs2791398	5.0E-04	9.8E-02	9.5E-02										
2	627,597	rs13388043	3.7E-03	1.2E-02	9.8E-04	1,831,232									
2	2,458,829	rs11127353	4.0E-03	7.3E-04	8.6E-02	276,395									
2	2,735,224	rs4853775	5.9E-07	8.5E-06	3.2E-04	256,099	Fail					Intergenic			Intergenic
2	2,991,323	rs17040704	6.7E-04	3.8E-02	1.7E-02	460,042									
2	3,451,365	rs12994580	2.1E-02	8.8E-04	3.0E-02	4,652,034									
2	8,103,399	rs6732658	6.0E-02	6.3E-04	6.2E-02	3,344,421									
2	11,447,820	rs4669724	4.7E-03	2.4E-02	9.1E-04	262,488									
2	11,710,308	rs11690767	1.1E-01	4.0E-02	8.9E-04	1,373,512									
2	13,083,820	rs953006	2.8E-04	6.6E-03	1.0E-02	95,134									
2	13,178,954	rs10208114	5.3E-04	2.1E-03	4.4E-03	9,398									
2	13,188,352	rs2128503	4.3E-04	1.5E-03	3.3E-03	52,320									
2	13,240,672	rs6719067	4.0E-04	1.9E-03	4.7E-03	302,149									
2	13,542,821	rs480881	2.2E-05	3.3E-05	7.7E-03	744	Pass					Intergenic	744		Intergeneic
2	13,543,565	rs511773	4.8E-04	9.8E-04	2.6E-02	5,247							5,247		Intergeneic
2	13,548,812	rs11693523	6.1E-05	1.3E-04	4.4E-03	20,648	Pass					Intergenic	20,648		Intergeneic
2	13,569,460	rs10929862	7.8E-05	2.2E-04	5.1E-03	86	Pass					Intergenic	86		Intergeneic
2	13,569,546	rs11686208	7.9E-05	2.3E-04	6.8E-03	656,868	Pass					Intergenic			Intergeneic
2	14,226,414	rs1448965	5.6E-03	6.0E-04	3.0E-02	4,197							4,197		Intergeneic
2	14,230,611	rs10427320	1.4E-02	3.6E-04	6.2E-02	3,887							3,887		Intergeneic
2	14,234,498	rs1448972	5.1E-03	4.4E-04	2.5E-02	1,760,718									Intergeneic
2	15,995,216	rs10929400	8.0E-03	4.9E-04	1.6E-02	1,546,313									
2	17,541,529	rs368430	9.0E-02	6.7E-02	5.2E-05	34,648	Pass					VSNL1	34,648		VSNL1
2	17,576,177	rs446947	1.0E-01	7.7E-02	1.7E-04	25,503							25,503		VSNL1
2	17,601,680	rs426590	1.2E-01	9.4E-02	1.7E-04	5,407,687									VSNL1
2	23,009,367	rs7586848	4.3E-04	3.6E-03	1.1E-02	6,792,835									
2	29,802,202	rs2257135	4.0E-03	1.2E-02	2.6E-04	5,966									
2	29,808,168	rs2631956	4.1E-03	4.1E-02	6.1E-04	79,615									
2	29,887,783	rs11694445	1.2E-04	6.3E-04	3.2E-04	7,938									
2	29,895,721	rs10187657	4.3E-03	5.5E-04	1.2E-03	98,323									
2	29,994,044	rs4233753	3.8E-03	3.9E-03	2.2E-04	1,152,277									
2	31,146,321	rs17010651	1.3E-10	2.1E-04	2.0E-04	122,260	Fail					GALNT14			GALNT14
2	31,268,581	rs10192210	6.6E-04	8.7E-04	4.0E-03	5,007,239									
2	36,275,820	rs1167437	1.5E-04	9.9E-05	8.0E-03	5,351,829	Pass					Intergenic			Intergenic
2	41,627,649	rs11124822	6.6E-03	9.3E-03	1.0E-03	1,628,152									
2	43,255,801	rs4952974	7.0E-04	1.4E-03	2.7E-03	1,201,769									
2	44,457,570	rs786626	1.0E-03	1.9E-02	8.6E-04	50,839									
2	44,508,409	rs1067367	7.1E-04	9.1E-03	1.2E-03	5,711							5,711		C2orf34
2	44,514,120	rs1065786	7.1E-04	8.9E-03	1.7E-03	11,549							11,549		C2orf34
2	44,525,669	rs1067347	3.7E-04	6.0E-03	7.4E-04	67,707									C2orf34
2	44,593,376	rs11888605	9.5E-04	1.0E-02	4.6E-03	6,391									
2	44,599,767	rs3862996	4.8E-04	6.2E-03	1.1E-03	1,104,546									
2	45,704,313	rs7565326	1.5E-03	7.4E-04	4.0E-03	557,531									
2	46,261,844	rs2594491	8.7E-03	2.0E-03	5.4E-04	5,892,860									
2	52,154,704	rs1806739	1.7E-02	2.8E-01	7.6E-04	4,776,793									
2	56,931,497	rs6545601	1.4E-02	2.8E-02	6.8E-04	26,666							26,666		LOC647016
2	56,958,163	rs2870503	1.4E-02	3.4E-02	6.8E-04	15,057							15,057		LOC647016
2	56,973,220	rs4672171	4.0E-03	8.2E-03	3.1E-04	34,382							34,382		LOC647016
2	57,007,602	rs2865254	5.9E-04	1.4E-03	2.1E-04	605,017									LOC647016
2	57,612,619	rs820985	7.7E-05	8.8E-04	5.7E-03	41,490	Pass					Intergenic	41,490		Intergenic
2	57,654,109	rs11125718	4.0E-04	2.5E-03	2.4E-02	4,400,872									Intergenic
2	62,054,981	rs11885927	7.3E-03	3.7E-04	1.9E-03	63,080									
2	62,118,061	rs12614102	7.1E-03	3.4E-04	8.2E-03	2,980,808									
2	65,098,869	rs759458	1.9E-01	2.4E-04	3.1E-02	1,928,536									
2	67,027,405	rs17032787	6.8E-04	4.8E-03	3.1E-04	3,886,214									
2	70,913,619	rs17006436	4.4E-04	2.6E-02	5.1E-02	2,386,708									
2	73,300,327	rs7605115	1.8E-04	2.3E-04	1.1E-04	3,674,397									
2	76,974,724	rs17013302	4.3E-04	8.9E-03	1.7E-01	222									
2	76,974,946	rs17013306	5.5E-04	8.8E-03	2.5E-01	6,214,322									
2	83,189,268	rs10496291	4.6E-04	2.8E-03	2.3E-02	3,149,558									
2	86,338,826	rs17510512	5.8E-04	2.8E-03	8.1E-03	536,628									
2	86,875,454	rs13023213	4.8E-03	9.5E-04	3.7E-02	15,833,416									
2	102,708,870	rs17027825	1.3E-02	7.7E-04	1.7E-02	32,618							32,618		TMEM182
2	102,741,488	rs2732838	5.6E-03	4.7E-05	2.7E-03	6,761	Pass					TMEM182	6,761		TMEM182
2	102,748,249	rs2732841	3.9E-03	4.5E-05	1.4E-03	16,517	Pass					TMEM182	16,517		TMEM182
2	102,764,766	rs2732860	1.2E-03	1.1E-04	1.1E-03	16,497							16,497		TMEM182
2	102,781,263	rs2110665	8.4E-04	5.9E-05	1.7E-03	1,277,253	Pass					TMEM182			TMEM182
2	104,058,516	rs11123977	1.0E-02	4.5E-03	1.5E-05	21,878	Pass					LOC728838	21,878		LOC728838
2	104,080,394	rs2376057	3.6E-02	2.7E-02	1.7E-04	169							169		LOC728838
2	104,080,563	rs2889276	6.1E-02	8.5E-02	5.2E-04	728,917									LOC728838
2	104,809,480	rs4851722	5.0E-03	5.3E-04	1.4E-02	643,896									

2	105,453,376	rs6725019	9.6E-04	2.6E-03	1.9E-02	557,828									
2	106,011,204	rs17032192	6.9E-03	1.2E-04	1.3E-02	12,076,316									
2	118,087,520	rs2551684	3.6E-04	3.2E-03	1.6E-03	27,452									
2	118,114,972	rs2255565	3.6E-04	3.4E-03	1.8E-03	1,484,829									
2	119,599,801	rs7570552	6.0E-04	4.6E-03	3.8E-03	370,016									
2	119,969,817	rs920338	1.1E-02	7.9E-04	1.9E-01	1,690,010									
2	121,659,827	rs12467309	7.4E-02	1.0E-02	6.9E-04	1,044,232									
2	122,704,059	rs1432311	3.6E-03	4.4E-04	3.7E-03	9,943									
2	122,714,002	rs934390	6.4E-04	5.3E-04	8.6E-04	95,396									
2	122,809,398	rs1881893	1.5E-02	4.4E-04	2.9E-02	25,853							25,853		Intergenic
2	122,835,251	rs10496582	1.4E-02	3.5E-04	1.6E-02	132							132		Intergenic
2	122,835,383	rs10496583	4.2E-02	3.9E-04	3.3E-02	425,221									Intergenic
2	123,260,604	rs6754547	2.2E-03	4.2E-03	1.3E-04	16,562							16,562		Intergenic
2	123,277,166	rs294665	7.7E-03	1.3E-02	7.7E-05	15,297	Pass					Intergenic	15,297		Intergenic
2	123,292,463	rs296042	1.8E-03	2.6E-02	8.8E-05	1,263	Pass					Intergenic	1,263		Intergenic
2	123,293,726	rs296047	2.4E-03	3.0E-02	1.5E-04	9,607							9,607		Intergenic
2	123,303,333	rs12469942	2.3E-03	2.8E-02	1.4E-04	12,610							12,610		Intergenic
2	123,315,943	rs296029	5.6E-03	2.6E-02	9.1E-04	3,327							3,327		Intergenic
2	123,319,270	rs296034	1.2E-03	2.2E-02	5.3E-05	12,113	Pass					Intergenic	12,113		Intergenic
2	123,331,383	rs294643	1.4E-03	1.9E-02	4.5E-05	73,232	Pass					Intergenic			Intergenic
2	123,404,615	rs2139011	1.1E-02	2.4E-02	2.3E-04	1,610,215									
2	125,014,830	rs17011610	1.5E-04	8.8E-03	9.2E-03	19,863									
2	125,034,693	rs6760693	9.0E-04	5.2E-02	1.0E-01	2,701,232									
2	127,735,925	rs4150512	1.8E-04	3.2E-02	2.8E-02	5,433,000									
2	133,168,925	rs9287423	2.1E-04	4.4E-03	1.7E-01	7,469,645									
2	140,638,570	rs12328474	1.0E-04	1.9E-02	3.7E-03	1,714,360									
2	142,352,930	rs10496905	2.3E-03	1.8E-03	8.4E-04	1,253,289									
2	143,606,219	rs354714	1.7E-02	6.1E-04	2.2E-03	400,937									
2	144,007,156	rs4525633	1.3E-02	6.3E-04	5.0E-03	6,301,799									
2	150,308,955	rs10490076	1.5E-01	8.4E-04	6.8E-02	3,280									
2	150,312,235	rs17745212	1.4E-01	7.1E-04	6.3E-02	1,149,914									
2	151,462,149	rs10188669	5.4E-02	3.8E-03	7.2E-05	2,978,234	Pass					Intergenic			Intergenic
2	154,440,383	rs11893783	9.2E-04	9.2E-03	9.7E-02	2,434,635									
2	156,875,018	rs16840214	2.3E-02	4.7E-04	1.1E-01	2,825,927									
2	159,700,945	rs2288105	1.9E-04	2.9E-02	3.8E-02	3,720,901									
2	163,421,846	rs1001911	8.4E-03	3.5E-04	6.4E-03	277							277		KCNH7
2	163,422,123	rs1492990	7.0E-03	2.4E-04	7.4E-03	20,253							20,253		KCNH7
2	163,442,376	rs10203602	1.1E-03	2.4E-04	3.7E-03	7,596							7,596		KCNH7
2	163,449,972	rs1364562	1.5E-03	3.0E-04	5.3E-03	19,865							19,865		KCNH7
2	163,469,837	rs307869	4.3E-03	6.4E-04	5.7E-03	9,912,552									KCNH7
2	173,382,389	rs7584826	2.7E-02	4.9E-04	4.3E-03	5,126,972									
2	178,509,361	rs13386418	1.8E-04	2.8E-03	4.5E-03	2,130,345									
2	180,639,706	rs262302	1.3E-04	6.3E-04	3.4E-03	8,682,875									
2	189,322,581	rs1400945	1.0E-02	9.7E-04	9.9E-03	1,691,488									
2	191,014,069	rs7581762	1.3E-01	1.3E-01	5.6E-04	213,036									
2	191,227,105	rs1860849	2.7E-02	1.3E-03	7.0E-04	1,533,493									
2	192,760,598	rs1113932	9.7E-02	9.6E-04	3.9E-02	8,250,657									
2	201,011,255	rs2028201	4.3E-03	8.7E-04	3.2E-03	3,417,129									
2	204,428,384	rs11571292	2.5E-02	4.0E-03	2.6E-04	20,727		CTLA4	204,440,754	204,446,928	Mixed	South Moroccan, UK, Russian, Filipii	20,727	CTLA4	CTLA4
2	204,449,111	rs231726	7.3E-03	3.4E-04	5.2E-04	300		CTLA4	204,440,754	204,446,928	Mixed	South Moroccan, UK, Russian, Filipii	300	CTLA4	CTLA4
2	204,449,411	rs1427676	6.5E-03	4.5E-04	6.0E-04	448,253		CTLA4	204,440,754	204,446,928	Mixed	South Moroccan, UK, Russian, Filipinos, Alsacian, DAISY, Ja		CTLA4	CTLA4
2	204,897,664	rs4673296	2.5E-03	5.2E-04	2.2E-03	7,975,373									
2	212,873,037	rs17418814	6.9E-04	5.7E-04	5.3E-04	8,380									
2	212,881,417	rs16848312	1.3E-03	8.3E-04	1.4E-03	2,458,130									
2	215,339,547	rs7599060	4.9E-03	4.8E-04	1.3E-02	855,437		BARD1	215,301,522	215,382,611					BARD1
2	216,194,984	rs6752170	6.2E-03	3.2E-04	5.4E-03	352,192									
2	216,547,176	rs11692400	5.9E-04	4.5E-03	1.4E-02	889,870									
2	217,437,046	rs1468860	1.4E-02	6.1E-02	3.7E-04	965,495									
2	218,402,541	rs6729330	3.9E-04	1.9E-02	1.6E-01	1,940,212									
2	220,342,753	rs1035680	1.5E-13	1.2E-10	1.5E-05	1,889,933	poor					Intergenic			Intergenic
2	222,232,686	rs13030635	1.0E-01	3.2E-02	7.8E-04	1,704,316									
2	223,937,002	rs1250915	1.7E-02	6.2E-04	8.4E-03	3,638									
2	223,940,640	rs990297	2.2E-02	7.6E-04	9.1E-03	726,918									
2	224,667,558	rs10197404	3.0E-02	5.7E-04	1.7E-02	175									
2	224,667,733	rs7574130	2.9E-02	4.7E-04	1.7E-02	3,545,014									
2	228,212,747	rs10179572	4.5E-02	7.8E-03	4.8E-04	1,701,272									
2	229,914,019	rs6757593	1.5E-04	1.2E-03	7.4E-03	3,262							3,262		DNER
2	229,917,281	rs7563766	9.7E-04	3.4E-03	1.4E-02	1,484							1,484		DNER
2	229,918,765	rs6717477	7.5E-04	3.3E-03	5.2E-03	4,430							4,430		DNER
2	229,923,195	rs9288646	2.3E-04	1.4E-03	4.3E-03	1,602							1,602		DNER
2	229,924,797	rs6748446	2.7E-04	1.6E-03	1.2E-02	7,478							7,478		DNER
2	229,932,275	rs7594321	3.7E-04	2.9E-03	9.5E-03	438							438		DNER
2	229,932,713	rs6436860	6.8E-04	5.0E-03	1.7E-02	32,117							32,117		DNER
2	229,964,830	rs207671	1.2E-04	5.6E-04	3.9E-03	83,812									DNER
2	230,048,642	rs6715897	3.2E-03	4.0E-04	2.5E-02	734							734		DNER
2	230,049,376	rs6750269	2.1E-03	2.2E-04	1.7E-02	11,586							11,586		DNER
2	230,060,962	rs7570424	1.8E-03	2.0E-04	1.7E-02	112							112		DNER
2	230,061,074	rs7609541	2.1E-03	2.5E-04	1.4E-02	21							21		DNER
2	230,061,095	rs7556846	1.8E-03	1.9E-04	1.7E-02	125							125		DNER
2	230,061,220	rs7570739	1.3E-03	1.3E-04	1.2E-02	5,229,679									DNER
2	235,290,899	rs4663163	3.5E-03	2.4E-04	8.5E-02	156									

2	235,291,055	rs4663165	1.8E-03	1.5E-04	6.9E-02	2,401,323							
2	237,692,378	rs7570534	1.3E-03	1.4E-04	3.5E-03	26,728						26,728	COPS8
2	237,719,106	rs6761680	2.8E-03	5.4E-04	1.3E-02	514						514	COPS8
2	237,719,620	rs7585219	7.2E-04	3.5E-04	3.3E-03	1,818						1,818	COPS8
2	237,721,438	rs10201392	6.3E-04	3.8E-04	3.3E-03	8,764						8,764	COPS8
2	237,730,202	rs10172321	1.5E-03	5.3E-04	3.9E-03	1,197,798							COPS8
2	238,928,000	rs821804	1.7E-05	7.1E-05	4.4E-05	135,548	Pass				TRAF3IP1		TRAF3IP1
2	239,063,548	rs523437	9.5E-05	1.7E-03	3.0E-04	1,621	Pass				TRAF3IP1	1,621	TRAF3IP1
2	239,065,169	rs580209	1.3E-04	2.0E-03	3.7E-04	1,490,800							TRAF3IP1
2	240,555,969	rs4854067	1.6E-03	2.5E-02	1.2E-04	28,161						28,161	NDUFA10
2	240,584,130	rs12328143	5.6E-04	2.5E-03	6.6E-04	3,078						3,078	NDUFA10
2	240,587,208	rs11685586	1.9E-03	3.0E-02	2.4E-04								NDUFA10
3	3,030,247	rs17024684	1.5E-02	2.9E-04	1.0E-02	5,666		CNTN4	2,117,247	3,074,645			CNTN4
3	3,035,913	rs1420020	3.1E-02	7.5E-04	1.8E-02	1,843,015		CNTN4	2,117,247	3,074,645			CNTN4
3	4,878,928	rs9819062	9.5E-04	9.6E-03	1.2E-02	469						469	ITPR1
3	4,879,397	rs12638937	1.0E-03	1.0E-02	8.7E-03	1,659						1,659	ITPR1
3	4,881,056	rs17041614	4.8E-04	5.6E-03	8.7E-03	1,188,317							ITPR1
3	6,069,373	rs1607193	1.3E-03	3.1E-02	1.7E-04	6,811,093							
3	12,880,466	rs11128620	1.0E-03	2.4E-02	1.0E-01	1,715,346							
3	14,595,812	rs9821893	2.8E-02	5.6E-04	2.8E-03	3,715							
3	14,599,527	rs11712736	6.8E-02	7.7E-04	4.7E-03	724,602							
3	15,324,129	rs10510436	4.1E-03	6.6E-04	5.3E-04	820,585							
3	16,144,714	rs17272271	5.1E-03	3.3E-04	5.2E-03	23,807							
3	16,168,521	rs2657589	4.8E-03	4.3E-04	9.9E-03	3,141,178							
3	19,309,699	rs727139	3.6E-05	9.5E-05	3.2E-04	8,579,831	poor				KCNH8		KCNH8
3	27,889,530	rs9849464	2.1E-04	1.3E-03	2.1E-02	39,511							
3	27,929,041	rs17020606	5.9E-04	9.4E-04	2.2E-02	1,164,299							
3	29,093,340	rs541374	5.4E-02	7.6E-04	1.7E-02	15,342,494							
3	44,435,834	rs9865920	2.6E-03	1.6E-04	1.5E-02	25,084						25,084	ZNF445
3	44,460,918	rs4610267	3.9E-03	2.1E-04	3.9E-02	11,687						11,687	ZNF445
3	44,472,605	rs9311357	4.8E-03	2.9E-04	3.8E-02	101,502							ZNF445
3	44,574,107	rs2293600	7.3E-03	4.5E-04	1.0E-01	11,355							
3	44,585,462	rs9867336	1.6E-02	4.8E-04	8.8E-02	60,759							
3	44,646,221	rs6769283	9.7E-03	6.4E-04	8.3E-02	13						13	ZNF197
3	44,646,234	rs6769288	9.8E-03	5.3E-04	9.4E-02	3,434						3,434	ZNF197
3	44,649,668	rs11710959	1.1E-02	6.8E-04	9.4E-02	69,329							ZNF197
3	44,718,997	rs6778410	4.2E-03	5.5E-04	6.6E-02	512,198							
3	45,231,195	rs28294	7.4E-02	5.7E-03	9.8E-04	277,033							
3	45,508,228	rs2128361	1.5E-07	2.8E-05	6.3E-04	15,021,333	poor				LARS2		LARS2
3	60,529,561	rs2594259	8.6E-04	1.9E-03	5.5E-03	234,267							
3	60,763,828	rs1447926	3.3E-03	3.1E-04	4.0E-03	1,105,912							
3	61,869,740	rs11919255	5.1E-02	3.0E-04	9.8E-03	4,532,012							
3	66,401,752	rs13088067	1.6E-03	4.1E-03	4.5E-04	400							
3	66,402,152	rs13097619	1.1E-03	3.0E-03	6.6E-04	7,548,983							
3	73,951,135	rs6784023	4.8E-04	8.1E-03	4.3E-03	3,454,203							
3	77,405,338	rs6784671	6.9E-03	8.8E-04	1.3E-01	4,962,595							
3	82,367,933	rs4364188	9.1E-05	8.1E-03	2.5E-02	38,068	Pass				Intergenic	38,068	Intergeneic
3	82,406,001	rs7645576	1.1E-04	1.1E-02	2.4E-02	13,783						13,783	Intergeneic
3	82,419,784	rs9309889	3.8E-04	2.5E-02	3.9E-02	4,619						4,619	Intergeneic
3	82,424,403	rs6778876	8.6E-04	3.1E-02	4.3E-02	9,792						9,792	Intergeneic
3	82,434,195	rs9878507	3.2E-05	2.7E-03	1.5E-02	7,018	Pass				Intergenic	7,018	Intergeneic
3	82,441,213	rs2122356	1.8E-04	1.3E-02	2.8E-02	512						512	Intergeneic
3	82,441,725	rs1449862	7.9E-05	8.4E-03	1.5E-02	10,421	Pass				Intergenic	10,421	Intergeneic
3	82,452,146	rs9880194	1.2E-04	8.8E-03	2.0E-02	19,230						19,230	Intergeneic
3	82,471,376	rs6789603	1.2E-04	1.6E-02	1.5E-02	14,106						14,106	Intergeneic
3	82,485,482	rs11916850	1.8E-04	1.0E-02	1.7E-02	10,694						10,694	Intergeneic
3	82,496,176	rs1449865	4.8E-05	4.2E-03	1.2E-02	15,664	Pass				Intergenic	15,664	Intergeneic
3	82,511,840	rs935725	3.8E-05	5.7E-03	1.8E-02	2,709	Pass				Intergenic	2,709	Intergeneic
3	82,514,549	rs9827304	8.7E-05	8.0E-03	3.2E-02	6,275	Pass				Intergenic	6,275	Intergeneic
3	82,520,824	rs9868998	3.3E-05	4.7E-03	1.5E-02	1,292	Pass				Intergenic	1,292	Intergeneic
3	82,522,116	rs12497206	1.7E-04	1.1E-02	2.0E-02	833						833	Intergeneic
3	82,522,949	rs2044752	3.6E-04	2.5E-02	3.3E-02	1,490						1,490	Intergeneic
3	82,524,439	rs9309900	4.1E-05	2.9E-03	9.6E-03	4,764	Pass				Intergenic	4,764	Intergeneic
3	82,529,203	rs6783449	9.1E-05	3.5E-03	1.0E-02	233	Pass				Intergenic	233	Intergeneic
3	82,529,436	rs2122355	1.7E-05	1.8E-03	7.1E-03	3,104,657	Pass				Intergenic		Intergeneic
3	85,634,093	rs726610	4.1E-02	8.2E-04	2.8E-02	39,705						39,705	Intergeneic
3	85,673,798	rs17517121	1.1E-01	9.7E-04	1.9E-02	12,535						12,535	Intergeneic
3	85,686,333	rs7652808	1.6E-02	2.6E-04	1.2E-02	6,844						6,844	Intergeneic
3	85,693,177	rs1530740	8.1E-02	6.6E-04	3.0E-02	12,176						12,176	Intergeneic
3	85,705,353	rs12494446	1.2E-01	9.9E-04	3.0E-02	3,273,585							Intergeneic
3	88,978,938	rs12634593	9.6E-04	1.4E-01	1.6E-02	33,570							
3	89,012,508	rs6764404	8.9E-04	1.2E-01	2.4E-02	97,838							
3	89,110,346	rs9829987	9.1E-04	2.9E-02	9.1E-03	13,353,193							
3	102,463,539	CG03018150	9.0E-04	9.1E-04	6.8E-02	15,821						15,821	
3	102,479,360	rs7634592	9.4E-04	1.7E-03	1.1E-01	24,977		SENP7	102,525,812	102,714,776		24,977	SENP7
3	102,504,337	rs16843699	8.5E-04	7.5E-04	6.8E-02	428,086		SENP7	102,525,812	102,714,776			SENP7
3	102,932,423	rs2611836	1.9E-03	9.5E-04	7.7E-02	3,092,789							
3	106,025,212	rs6437543	2.8E-03	1.5E-04	2.0E-02	129,729							
3	106,154,941	rs1369905	8.0E-03	2.6E-04	4.3E-03	8,224,122							
3	114,379,063	rs12495201	7.7E-04	2.8E-02	1.8E-01	17,764,610							
3	132,143,673	rs2296024	8.4E-04	3.2E-02	2.0E-01	9,532,419							

3	141,676,092	rs10513115	5.9E-04	1.6E-01	3.1E-01	5,231,644								
3	146,907,736	rs2459136	8.7E-05	5.4E-03	1.5E-03	18,363	Pass				GM2AP	18,363		GM2AP
3	146,926,099	rs2687861	2.1E-04	1.2E-02	6.6E-03	4,918,331								GM2AP
3	151,844,430	rs6440690	8.6E-05	3.6E-02	4.4E-02	785,991	Pass				SELT			SELT
3	152,630,421	rs6771904	1.8E-02	5.3E-04	2.7E-03	1,137,684		MED12L	152,287,366	152,634,500				MED12L
3	153,768,105	rs11925597	4.9E-02	5.5E-04	1.6E-01	4,576,339								
3	158,344,444	rs6809003	2.7E-03	9.1E-05	7.1E-03	88,982	Pass				CCNL1			CCNL1
3	158,433,426	rs13081038	4.6E-03	1.9E-04	1.2E-02	1,462						1,462		CCNL1
3	158,434,888	rs10451916	3.1E-03	1.5E-04	1.1E-02	551						551		CCNL1
3	158,435,439	rs1392799	1.2E-03	2.1E-05	4.3E-03	10,853	Pass				CCNL1	10,853		CCNL1
3	158,446,292	rs9831830	3.0E-03	1.6E-04	1.1E-02	2,637						2,637		VEPH1
3	158,448,929	rs7642179	3.0E-03	1.0E-04	6.3E-03	7,569						7,569		VEPH1
3	158,456,498	rs1500925	3.4E-03	1.8E-04	1.1E-02	627						627		VEPH1
3	158,457,125	rs9289980	1.5E-02	7.1E-04	1.2E-02	7,054						7,054		VEPH1
3	158,464,179	rs924517	1.5E-02	8.9E-04	8.4E-03	9,907						9,907		VEPH1
3	158,474,086	rs12633429	2.5E-02	9.6E-04	8.9E-03	17,521						17,521		VEPH1
3	158,491,607	rs2874616	2.7E-02	9.6E-04	1.3E-02	35,260						35,260		VEPH1
3	158,526,867	rs10936081	8.7E-03	2.8E-04	9.7E-03	1,677						1,677		VEPH1
3	158,528,544	rs1522392	2.3E-02	8.5E-04	2.1E-02	293,539								VEPH1
3	158,822,083	rs11709879	1.7E-02	7.6E-03	5.8E-04	2,018								
3	158,824,101	rs9876844	1.7E-02	9.2E-03	5.6E-04	1,657,233								
3	160,481,334	rs9831067	5.4E-02	6.1E-04	1.6E-02	194,103								
3	160,675,437	rs9824310	1.7E-02	3.4E-03	9.4E-04	1,799,688								
3	162,475,125	rs9828189	3.9E-02	7.9E-02	2.2E-04	37,361						37,361		NMD3
3	162,512,486	rs9811833	3.6E-02	7.9E-02	2.2E-04	29,825						29,825		NMD3
3	162,542,311	rs336546	3.0E-02	7.6E-02	2.5E-04	5,683,665								NMD3
3	168,225,976	rs1949532	1.7E-05	9.4E-05	1.3E-04	216,996	Pass				Intergenic			Intergenic
3	168,442,972	rs11923256	4.1E-04	1.2E-02	8.0E-03	141,846								
3	168,584,818	rs1554568	3.6E-04	3.1E-03	2.5E-02	1,487						1,487		SERPINI2
3	168,586,305	rs11714970	6.9E-04	5.0E-03	2.9E-02	7,778						7,778		SERPINI2
3	168,594,083	CG03028974	9.9E-04	8.4E-03	2.8E-02	8,944,150								SERPINI2
3	177,538,233	rs2128850	4.1E-04	8.7E-02	2.8E-02	659						659		LOC730168
3	177,538,892	rs13100792	8.1E-04	1.4E-01	2.4E-02	22,262						22,262		LOC730168
3	177,561,154	rs2170808	9.0E-04	1.2E-01	7.6E-02	8,056,977								LOC730168
3	185,618,131	rs1533678	3.2E-04	4.8E-02	1.2E-01	19,971								
3	185,638,102	rs7429010	9.8E-04	5.1E-02	2.2E-01	200,280								
3	185,838,382	rs7646910	6.1E-03	3.0E-02	8.0E-04	667								
3	185,839,049	rs4686705	8.4E-03	3.8E-02	8.0E-04	1,445,273								
3	187,284,322	rs7433760	5.9E-02	9.1E-04	2.6E-03	4,365						4,365		ETV5
3	187,288,687	rs4686730	6.4E-02	1.9E-04	1.3E-03	5,502						5,502		ETV5
3	187,294,189	rs6444106	7.0E-02	2.8E-04	1.4E-03	303,984								ETV5
3	187,598,173	rs9851300	2.4E-02	1.9E-02	5.4E-04	1,677,592								
3	189,275,765	rs9813167	6.5E-04	3.1E-04	1.5E-03	458,465								
3	189,734,230	rs17605260	1.2E-03	1.7E-02	8.8E-04	9,012						9,012		LPP
3	189,743,242	rs1426261	5.0E-04	1.1E-02	3.1E-04	155						155		LPP
3	189,743,397	rs1426260	1.0E-03	1.2E-02	6.2E-04	7,909						7,909		LPP
3	189,751,306	rs4263285	1.1E-03	2.1E-02	3.4E-04	1,716,531								LPP
3	191,467,837	rs17427594	2.3E-01	7.0E-02	5.0E-04	4,543,255								
3	196,011,092	rs4434099	1.9E-03	9.3E-04	1.9E-03	1,380,958								
3	197,392,050	rs4916483	3.7E-02	5.9E-04	7.2E-02									
4	5,820,851	rs6817935	5.4E-03	6.1E-04	1.6E-01	205,170								
4	6,026,021	rs11733383	1.0E-01	8.0E-04	4.9E-02	1,712,997								
4	7,739,018	rs1557818	2.4E-03	3.6E-04	7.1E-04	3,075								
4	7,742,093	rs929264	4.6E-04	1.8E-04	4.0E-04	847,788								
4	8,589,881	rs3103091	4.4E-04	8.3E-02	1.0E-01	7,458,654								
4	16,048,535	rs4698487	1.3E-04	3.7E-03	1.1E-03	284,523								
4	16,333,058	rs1514645	6.9E-04	7.6E-05	6.5E-04	2,246	Pass				LDB2	2,246		LDB2
4	16,335,304	rs2645255	5.8E-04	9.3E-05	8.9E-04	580	Pass				LDB2	580		LDB2
4	16,335,884	rs2658509	2.2E-02	7.7E-04	6.6E-03	10,844						10,844		LDB2
4	16,346,728	rs283021	1.2E-02	6.1E-04	7.3E-03	6,454						6,454		LDB2
4	16,353,182	rs883086	5.3E-04	4.7E-04	1.4E-03	1,314,537								LDB2
4	17,667,719	rs9784574	1.6E-01	7.1E-03	9.7E-04	1,157,874								
4	18,825,593	rs10516333	1.3E-03	7.6E-04	8.5E-03	6,374,818								
4	25,200,411	rs939352	2.1E-04	6.2E-03	4.7E-02	3,863,633								
4	29,064,044	rs13118914	3.2E-03	7.2E-04	5.2E-02	3,508,968								
4	32,573,012	rs6531531	1.0E-05	1.7E-03	1.9E-04	3,408,030	Pass				Intergenic			Intergenic
4	35,981,042	rs10001947	4.7E-02	8.5E-04	4.8E-02	6,514,171								
4	42,495,213	rs2575515	8.2E-05	6.9E-03	1.1E-03	14,833	Pass				LOC389207	14,833		LOC389207
4	42,510,046	rs4861224	7.8E-04	2.5E-02	3.6E-03	99,351								LOC389207
4	42,609,397	rs4861237	1.3E-03	1.9E-04	2.2E-02	25,052						25,052		LOC389207
4	42,634,449	rs2345759	6.5E-03	3.8E-04	1.0E-01	20,976						20,976		LOC389207
4	42,655,425	rs992133	4.0E-03	2.3E-04	6.4E-02	64,587								LOC389207
4	42,720,012	rs10938230	8.6E-03	5.2E-04	1.1E-01	17,151						17,151		LOC389207
4	42,737,163	rs920153	5.2E-03	3.7E-04	6.1E-02	6,839						6,839		LOC389207
4	42,744,002	rs10938235	7.2E-03	3.0E-04	9.8E-02	5,412						5,412		LOC389207
4	42,749,414	rs720961	9.0E-03	3.6E-04	1.1E-01	4,307						4,307		LOC389207
4	42,753,721	rs6855602	9.1E-03	3.6E-04	1.1E-01	7,773						7,773		LOC389207
4	42,761,494	rs1595187	7.2E-03	3.4E-04	1.3E-01	12,833,910								LOC389207
4	55,595,404	rs4864949	9.7E-04	1.3E-01	3.1E-02	339,640								
4	55,935,044	rs4864984	1.8E-02	3.6E-04	1.1E-02	12,537								
4	55,947,581	rs7664496	3.5E-02	8.3E-04	3.4E-02	5,144,943								

4	61,092,524	rs1511121	1.1E-01	6.7E-04	1.3E-01	17,800								
4	61,110,324	rs17090188	2.7E-02	9.1E-04	3.1E-02	471,272								
4	61,581,596	rs6829909	2.7E-04	3.2E-03	3.3E-03	4,485,359								
4	66,066,955	rs4604104	6.2E-04	2.5E-02	2.5E-02	7,892,583								
4	73,959,538	rs186403	5.0E-05	7.9E-03	2.1E-04	29,704	Pass				LOC728019	29,704		LOC728019
4	73,989,242	rs919047	7.2E-04	4.7E-02	7.2E-04	14,901,720								LOC728019
4	88,890,962	rs11733008	2.0E-02	3.9E-02	4.8E-04	4,970,186								
4	93,861,148	rs10155056	9.7E-04	1.6E-03	1.8E-03	64,919								
4	93,926,067	rs17259489	7.1E-03	4.1E-03	6.8E-04	265,258								
4	94,191,325	rs4491988	4.1E-03	5.5E-04	4.6E-04	634								
4	94,191,959	rs2870656	1.7E-03	5.3E-04	5.8E-04	3,190,528								
4	97,382,487	rs17025219	1.6E-04	4.4E-03	3.6E-03	1,418						1,418		Intergenic
4	97,383,905	rs7653937	1.6E-04	4.1E-03	5.4E-03	24,635						24,635		Intergenic
4	97,408,540	rs5017329	1.0E-04	3.6E-03	4.4E-03	10,351	Pass					10,351		Intergenic
4	97,418,891	rs6830746	1.4E-04	3.0E-03	2.8E-03	41,669						41,669		Intergenic
4	97,460,560	rs2166016	1.2E-04	3.7E-03	4.4E-03	18,818						18,818		Intergenic
4	97,479,378	rs1605227	3.0E-04	7.5E-03	7.8E-03	5,946						5,946		Intergenic
4	97,485,324	rs6832934	1.9E-04	4.7E-03	6.6E-03	109,981								
4	97,595,305	rs4699289	8.2E-04	1.8E-02	1.2E-02	9,423,460								
4	107,018,765	rs10516529	1.2E-03	6.7E-04	4.9E-02	2,395,045								
4	109,413,810	rs2880544	5.7E-03	1.6E-03	2.9E-04	62,495								
4	109,476,305	rs17038826	5.1E-03	5.1E-04	2.3E-04	14,921								
4	109,491,226	rs917412	8.2E-03	8.7E-04	1.1E-05	1,243,784	Pass				Intergenic			Intergenic
4	110,735,010	rs6851009	3.9E-02	7.1E-04	9.2E-03	798,009								
4	111,533,019	rs2348205	7.0E-03	2.6E-03	9.8E-04	3,190,748								
4	114,723,767	rs10488894	7.2E-05	4.8E-02	4.1E-02	6,063,316	Pass				CAMK2D			CAMK2D
4	120,787,083	rs2389883	6.4E-02	8.3E-03	2.7E-04	372,050								
4	121,159,133	rs1503534	1.1E-02	8.2E-04	1.7E-01	504,018								
4	121,663,151	rs7441640	2.3E-03	9.4E-04	1.4E-03	179,815								
4	121,842,966	rs1496276	8.6E-04	2.4E-03	4.6E-02	1,394,618								
4	123,237,584	rs3108397	9.4E-04	2.1E-01	1.2E-01	8,780,187								
4	132,017,771	rs17502137	2.4E-03	7.9E-03	4.0E-05	2,915,886	Pass				Intergenic			Intergenic
4	134,933,657	CG04022664	5.6E-02	2.6E-04	1.7E-02	2,304,178								
4	137,237,835	rs1344033	2.7E-04	3.2E-03	3.7E-03	3,527,821								
4	140,765,656	rs11945690	5.1E-04	5.9E-02	4.5E-02	10,635,063								
4	151,400,719	rs10003215	2.6E-06	3.6E-04	3.3E-04	3,612,222	Fail				LRBA			LRBA
4	155,012,941	rs1456385	3.9E-04	1.6E-02	1.7E-02	6,306						6,306		SFRP2
4	155,019,247	rs10517581	3.3E-04	1.5E-02	1.8E-02	145						145		SFRP2
4	155,019,392	CG04025829	4.2E-04	1.5E-02	1.8E-02	24,389						24,389		SFRP2
4	155,043,781	rs17299833	1.6E-04	7.1E-03	9.6E-03	7,858,154								SFRP2
4	162,901,935	rs1560925	3.8E-03	1.4E-02	9.0E-04	2,331,689								
4	165,233,624	rs7659197	1.7E-02	3.6E-04	2.4E-01	4,856								
4	165,238,480	rs6536795	7.7E-03	2.6E-04	1.2E-01	3,521,881								
4	168,760,361	rs7657818	3.1E-04	1.4E-03	1.2E-02	1,082,132								
4	169,842,493	rs4692947	1.1E-04	3.1E-04	5.6E-04	958,398		PALLD	169,654,792	170,086,183				PALLD
4	170,800,891	rs17627781	1.5E-06	2.4E-05	3.4E-05	1,572,837	Fail				CLCN3			CLCN3
4	172,373,728	rs1038819	6.3E-03	7.7E-03	4.5E-04	5,332,245								
4	177,705,973	rs12511856	9.9E-04	4.8E-05	3.0E-02	648,458	Pass				Intergenic			Intergenic
4	178,354,431	rs10446826	8.1E-02	3.3E-02	4.5E-04	35,439								
4	178,389,870	rs6821871	1.1E-01	4.6E-03	8.3E-04	4,078,606								
4	182,468,476	rs10520475	8.9E-04	9.1E-04	2.5E-02	1,122,957								
4	183,591,433	rs11132124	4.0E-02	2.3E-03	7.3E-04	2,574,340								
4	186,165,773	rs11737460	7.3E-21	4.8E-14	2.1E-10	2,872,754	Fail				HELT			HELT
4	189,038,527	rs4862838	1.0E-04	1.0E-02	6.9E-02	1,031,117								
4	190,069,644	rs11737653	3.0E-04	9.5E-03	3.4E-02									
5	3,332,885	rs4645337	5.1E-04	6.7E-03	1.1E-02	3,222,269								
5	6,555,154	rs6873584	1.0E-02	9.4E-04	1.0E-02	394,103								
5	6,949,257	rs903867	2.4E-03	1.5E-02	8.7E-05	11,635	Pass				Intergenic	11,635		Intergenic
5	6,960,892	rs903863	1.3E-03	1.1E-02	4.7E-05	27,095	Pass				Intergenic	27,095		Intergenic
5	6,987,987	rs11134194	3.5E-03	4.0E-02	2.3E-04	3,096,547								Intergenic
5	10,084,534	rs10062164	5.6E-04	1.3E-03	1.3E-02	3,362,718								
5	13,447,252	rs985766	2.1E-03	5.6E-03	4.7E-04	2,032						2,032		Intergenic
5	13,449,284	rs6875759	2.9E-03	6.4E-03	4.3E-04	29,820						29,820		Intergenic
5	13,479,104	rs4702943	2.7E-03	7.2E-03	4.0E-04	12,354						12,354		Intergenic
5	13,491,458	rs10064110	2.2E-03	9.2E-03	4.1E-04	499,018								Intergenic
5	13,990,476	rs17278234	3.9E-04	2.9E-03	2.6E-03	1,035,547								
5	15,026,023	rs835087	1.3E-04	8.9E-03	1.9E-02	10,600						10,600		LOC402198
5	15,036,623	rs835071	1.6E-05	2.7E-03	3.0E-02	120	Pass				LOC402198	120		LOC402198
5	15,036,743	rs835070	5.9E-04	2.4E-02	8.8E-02	2,783,566								LOC402198
5	17,820,309	rs10941538	8.0E-04	8.3E-02	3.0E-01	106,574								
5	17,926,883	rs1542128	8.8E-04	6.3E-02	1.3E-01	225,841								
5	18,152,724	rs16887248	1.5E-04	4.2E-02	3.8E-03	417								
5	18,153,141	rs17660815	2.1E-04	5.7E-02	7.6E-03	1,376,300								
5	19,529,441	rs1486849	2.4E-02	7.6E-04	2.2E-01	933						933		CDH18
5	19,530,374	rs17216535	1.1E-02	6.7E-04	9.6E-02	781						781		CDH18
5	19,531,155	rs1486851	1.1E-02	7.2E-04	1.3E-01	30,693						30,693		CDH18
5	19,561,848	rs3811865	8.4E-03	8.2E-04	1.5E-01	5,702,136								CDH18
5	25,263,984	rs3913960	7.2E-03	6.9E-04	6.9E-04	7,397,253								
5	32,661,237	rs864674	3.1E-03	7.1E-04	2.5E-03	4,102,149								
5	36,763,386	rs36650	7.9E-03	6.7E-04	1.3E-02	17,290,603								
5	54,053,989	rs3886214	7.1E-03	3.0E-02	2.7E-04	470,687								

5	54,524,676	rs12110186	2.0E-02	1.6E-04	2.4E-02	863,851					
5	55,388,527	rs10042443	2.2E-04	1.4E-05	1.4E-04	2,299	Pass	IL6ST	2,299	IL6ST	
5	55,390,826	rs10044080	1.1E-03	3.9E-05	1.2E-03	502,606	Pass	IL6ST		IL6ST	
5	55,893,432	rs4301182	2.8E-03	7.4E-05	7.9E-03	4,995,848	Pass	Intergenic		Intergenic	
5	60,889,280	rs13177031	5.1E-05	1.7E-03	2.7E-03	6,453,252	Pass	ZSWIM6		ZSWIM6	
5	67,342,532	rs281441	4.0E-02	1.6E-04	2.3E-02	10,759,525					
5	78,102,057	rs10068924	6.3E-02	3.1E-02	5.8E-04	3,930,423					
5	82,032,480	rs1827389	1.2E-01	4.3E-04	1.8E-02	7,279,762					
5	89,312,242	rs10474312	4.0E-03	1.4E-04	7.5E-04	9,768					
5	89,322,010	rs1560208	2.4E-03	1.1E-04	2.1E-04	78,065					
5	89,400,075	rs818944	3.9E-03	2.5E-04	8.8E-04	10,846			10,846	Intergenic	
5	89,410,921	rs256515	3.1E-03	2.6E-04	4.9E-04	373			373	Intergenic	
5	89,411,294	rs181930	4.5E-03	2.8E-04	7.1E-04	13,721			13,721	Intergenic	
5	89,425,015	rs157559	2.1E-03	1.1E-04	2.0E-04	15,643			15,643	Intergenic	
5	89,440,658	rs924628	3.5E-03	3.5E-04	3.4E-04	10,555,286				Intergenic	
5	99,995,944	rs32282	3.6E-02	6.3E-03	8.0E-04	164,532					
5	100,160,476	rs32477	2.5E-01	3.7E-03	5.7E-04	9,199,107					
5	109,359,583	rs36789	5.1E-04	2.7E-02	1.1E-01	52,622					
5	109,412,205	rs17162591	4.3E-04	3.7E-03	2.6E-03	1,648,803					
5	111,061,008	rs17133647	8.0E-04	4.2E-02	7.1E-01	874,849					
5	111,935,857	rs9790990	9.7E-04	1.8E-02	1.0E-02	6,760,092					
5	118,695,949	rs986649	1.9E-02	2.6E-03	5.6E-04	328,106					
5	119,024,055	rs780428	6.2E-02	1.0E-03	6.4E-03	51,671					
5	119,075,726	rs11241515	1.9E-02	9.2E-04	3.3E-02	1,866,497					
5	120,942,223	rs4307116	7.3E-03	1.9E-04	5.8E-02	1,563,550					
5	122,505,773	rs467069	6.7E-03	7.8E-05	3.3E-03	6,015,793	Pass	PRDM6		PRDM6	
5	128,521,566	rs7726142	1.5E-01	1.0E-01	9.3E-04	6,451,415					
5	134,972,981	rs7715094	1.2E-02	1.2E-02	9.4E-04	8,261,520					
5	143,234,501	rs10075045	8.0E-04	5.5E-03	1.7E-02	10,838					
5	143,245,339	rs9324934	5.4E-04	6.7E-03	2.2E-02	89,330					
5	143,334,669	rs312622	1.6E-04	4.2E-02	2.9E-02	1,159,814					
5	144,494,483	rs7710868	2.0E-02	8.7E-04	2.0E-02	1,682,578					
5	146,177,061	rs1432892	2.7E-03	9.4E-04	5.6E-04	11,277,493					
5	157,454,554	rs1744344	1.9E-04	3.6E-02	7.1E-03	618,246					
5	158,072,800	rs3734105	2.6E-03	6.2E-04	2.4E-02	5,062			5,062	EBF1	
5	158,077,862	rs4147383	2.3E-03	8.3E-04	2.5E-02	49,129			49,129	EBF1	
5	158,126,991	rs10067813	7.8E-04	2.7E-03	1.8E-03	752,729				EBF1	
5	158,879,720	rs2421051	3.1E-02	1.0E-02	9.7E-04	151,381					
5	159,031,101	rs4616877	7.6E-02	6.5E-04	2.0E-02	147,200					
5	159,178,301	rs17057161	5.9E-04	2.3E-02	9.1E-05	9,150,496	Pass	Intergenic		Intergenic	
5	168,328,797	rs10053858	6.1E-04	3.7E-03	1.3E-03	206					
5	168,329,003	rs10053973	4.0E-04	1.1E-03	1.4E-04	446,983					
5	168,775,986	rs929758	2.2E-04	4.9E-05	8.3E-05	57	Pass	Intergenic	57	Intergenic	
5	168,776,043	rs929759	1.8E-04	5.4E-05	5.2E-05	19,967	Pass	Intergenic	19,967	Intergenic	
5	168,796,010	rs39797	2.8E-04	1.7E-04	1.6E-04	3,411,859				Intergenic	
5	172,207,869	rs792965	5.0E-06	4.1E-05	1.1E-04	2,339,895	Pass	ERGIC1		ERGIC1	
5	174,547,764	rs2381985	4.6E-02	2.5E-01	4.4E-05	131	Pass	Intergenic	131	Intergenic	
5	174,547,895	rs2381986	1.1E-01	4.2E-01	2.2E-04	14,338			14,338	Intergenic	
5	174,562,233	rs6556193	4.4E-02	5.6E-02	3.8E-04					Intergenic	
6	374,457	rs950286	5.2E-05	6.7E-06	4.8E-03	12,749	Pass	IRF4	12,749	IRF4	
6	387,206	rs4959880	3.8E-04	1.1E-04	1.0E-02	7,277			7,277	IRF4	
6	394,483	rs13214605	1.6E-03	8.7E-04	1.3E-01	569,823				IRF4	
6	964,306	rs845887	7.1E-02	2.9E-04	3.3E-02	2,165					
6	966,471	rs845882	1.0E-01	9.6E-04	3.0E-02	3,762,957					
6	4,729,428	rs17138843	6.8E-05	1.9E-02	1.3E-02	1,908,891	Pass	CDYL		CDYL	
6	6,638,319	rs2235493	6.6E-04	2.6E-02	2.2E-02	561,321					
6	7,199,640	rs2714342	1.4E-01	5.5E-02	4.4E-04	1,787,791					
6	8,987,431	rs9328516	4.3E-03	7.3E-05	5.4E-04	20,029	Pass	LOC389365	20,029	LOC389365	
6	9,007,460	rs2093745	1.5E-02	1.5E-04	2.1E-03	6,167			6,167	LOC389365	
6	9,013,627	rs2209123	1.3E-02	1.5E-04	2.2E-03	1,482,456				LOC389365	
6	10,496,083	rs4527693	1.4E-03	7.9E-04	6.7E-04	5,020,424					
6	15,516,507	rs2237150	2.8E-01	2.9E-01	9.3E-04	21,203			21,203	JARID2	
6	15,537,710	rs4085876	1.3E-01	4.1E-01	5.2E-04	7,659			7,659	JARID2	
6	15,545,369	rs2282827	8.9E-03	6.2E-02	5.1E-04	883,749				JARID2	
6	16,429,118	rs2038018	9.9E-04	1.3E-02	2.9E-02	688					
6	16,429,806	rs2237224	8.3E-04	9.5E-03	1.9E-02	3,441,826					
6	19,871,632	rs9350178	1.6E-04	2.4E-04	1.6E-04	1,331,085					
6	21,202,717	rs6939622	4.6E-02	3.5E-03	2.7E-04	2,847,139					
6	24,049,856	rs9466992	9.7E-07	3.5E-05	5.6E-04	3,831,955	Pass	LOC645385		LOC645385	
6	27,881,811	rs200481	4.4E-04	2.6E-04	9.2E-02	34,360			34,360	Histone Gene Cluster	
6	27,916,171	rs201002	1.6E-04	7.4E-05	4.9E-02	5,502	Pass	Histone Gene Cluster	5,502	Histone Gene Cluster	
6	27,921,673	rs200995	1.9E-04	7.4E-05	4.9E-02	1,800	Pass	Histone Gene Cluster	1,800	Histone Gene Cluster	
6	27,923,473	rs200991	2.6E-04	2.6E-04	4.3E-02	154,537				Histone Gene Cluster	
6	28,078,010	rs149946	1.8E-03	9.1E-04	5.1E-02	41,621			41,621	Olfactory Receptor	
6	28,119,631	rs202906	2.8E-04	1.0E-04	3.5E-02	9,617			9,617	Olfactory Receptor	
6	28,129,248	rs203887	9.8E-04	3.2E-04	2.6E-02	320			320	Olfactory Receptor	
6	28,129,568	rs203888	9.7E-05	9.1E-05	1.5E-02	2,162	Pass	Olfactory Receptor	2,162	Olfactory Receptor	
6	28,131,730	rs9393881	9.6E-04	6.1E-04	3.9E-02	183,612				Olfactory Receptor	
6	28,315,342	rs1150709	6.7E-04	6.9E-03	2.9E-01	100,671					
6	28,416,013	rs17301128	8.5E-04	2.3E-03	1.3E-01	1,102					
6	28,417,115	rs1119211	2.2E-04	5.5E-04	5.2E-02	224,810					

6	28,641,925	rs418092	9.3E-05	6.5E-05	2.8E-02	41,226	Pass				HLA Region		
6	28,683,151	rs1474986	1.8E-04	1.3E-03	2.9E-02	159,504					HLA Region		
6	28,842,655	rs1233604	1.0E-03	2.8E-04	5.0E-02	61,180					HLA Region		
6	28,903,835	rs3135296	8.0E-05	5.6E-05	1.7E-02	235,286	Pass				HLA Region		
6	29,139,121	rs3117143	1.8E-04	4.5E-05	2.7E-02	26,497	Pass				HLA Region		
6	29,165,618	rs3129788	6.0E-05	4.2E-05	2.3E-02	15,455	Pass				HLA Region		
6	29,181,073	rs3131091	3.1E-04	1.2E-03	9.8E-02	94,481					HLA Region		
6	29,275,554	rs3116830	3.4E-05	4.1E-05	1.2E-02	10,939	Pass				HLA Region		
6	29,286,493	rs2206040	5.2E-04	1.9E-03	1.2E-01	16,497					HLA Region		
6	29,302,990	rs3130813	7.8E-05	3.6E-04	8.7E-03	30,784	Pass				HLA Region		
6	29,333,774	rs3117330	7.5E-04	5.5E-03	6.3E-02	30,625					HLA Region		
6	29,364,399	rs1884123	9.3E-04	5.2E-03	4.2E-02	4,011					HLA Region		
6	29,368,410	rs3117425	2.0E-05	5.1E-06	2.9E-03	82,344	Pass				HLA Region		
6	29,450,754	rs3749971	7.2E-05	1.0E-05	7.7E-03	22,176	Pass				HLA Region		
6	29,472,930	rs2073151	7.8E-04	1.9E-03	5.9E-03	11,434					HLA Region		
6	29,484,364	rs406511	4.9E-06	2.8E-07	4.4E-04	39,079	Pass				HLA Region		
6	29,523,443	rs2523443	1.1E-05	6.3E-06	1.3E-03	28,657	Pass				HLA Region		
6	29,552,100	rs1233495	9.2E-03	8.2E-04	9.4E-02	102,678					HLA Region		
6	29,654,778	rs1233396	3.0E-04	1.4E-04	8.0E-03	1,290					HLA Region		
6	29,656,068	rs926552	9.4E-04	4.1E-04	3.2E-03	438,235					HLA Region		
6	30,094,303	rs3115631	6.6E-06	6.9E-06	2.6E-03	106,813	Pass				HLA Region		
6	30,201,116	rs2517592	8.8E-03	6.5E-04	1.9E-02	10,223					HLA Region		
6	30,211,339	rs1541269	5.9E-06	1.4E-06	4.5E-04	49,107	Pass				HLA Region		
6	30,260,446	CG06006774	1.4E-06	2.0E-06	4.4E-04	5,660	Pass				HLA Region		
6	30,266,106	rs2517618	5.9E-04	1.0E-04	5.9E-03	11,200					HLA Region		
6	30,277,306	rs2517611	1.8E-03	9.1E-04	6.2E-03	62,502					HLA Region		
6	30,339,808	rs3094070	1.1E-03	6.8E-04	2.4E-02	421					HLA Region		
6	30,340,229	rs3130404	9.3E-04	5.3E-04	2.0E-02	1,902					HLA Region		
6	30,342,131	rs3130405	7.7E-04	6.7E-04	2.9E-02	516					HLA Region		
6	30,342,647	rs2844764	1.6E-03	7.6E-04	1.6E-02	86,389					HLA Region		
6	30,429,036	rs3132649	6.4E-08	6.5E-06	2.1E-04	132	Pass				HLA Region		
6	30,429,168	rs3094061	3.6E-06	2.3E-05	1.2E-03	147	Pass				HLA Region		
6	30,429,315	rs3130374	7.2E-06	4.3E-05	8.6E-04	146	Pass				HLA Region		
6	30,429,461	rs3094627	5.0E-07	7.5E-06	5.2E-04	6,710	Pass				HLA Region		
6	30,436,171	rs3130351	1.2E-06	2.5E-06	1.8E-04	165	Pass				HLA Region		
6	30,436,336	rs3130352	1.1E-06	1.1E-06	1.2E-04	1,609	Pass				HLA Region		
6	30,437,945	rs3094057	4.8E-06	6.6E-06	2.6E-04	2,180	Pass				HLA Region		
6	30,440,125	rs3094055	4.3E-05	7.1E-04	2.1E-03	409	Pass				HLA Region		
6	30,440,534	rs1012411	1.1E-05	6.8E-05	1.1E-03	9,957	Pass				HLA Region		
6	30,450,491	rs3132636	9.4E-04	1.6E-03	3.9E-03	739					HLA Region		
6	30,451,230	rs3129819	7.2E-04	1.2E-03	4.5E-03	3,988					HLA Region		
6	30,455,218	rs970269	6.8E-04	9.5E-04	3.9E-03	1,962					HLA Region		
6	30,457,180	rs3094623	8.1E-04	9.8E-04	5.1E-03	9,390					HLA Region		
6	30,466,570	rs3094050	5.6E-11	7.1E-09	7.0E-06	366	Pass				HLA Region		
6	30,466,936	rs3094703	4.0E-10	2.1E-08	1.7E-05	4,394	Pass				HLA Region		
6	30,471,330	rs3094034	7.3E-07	1.5E-06	7.7E-05	88,553	Pass				HLA Region		
6	30,559,883	rs3094694	1.4E-07	1.3E-07	6.7E-05	176,178	Pass	HLA-E	30,565,250	30,569,077	HLA Region	HLA-E	
6	30,736,061	rs3130000	3.3E-05	1.3E-03	1.6E-03	65,734	Pass				HLA Region		
6	30,801,795	rs3095329	2.6E-08	7.3E-07	8.3E-06	3,631	Pass				HLA Region		
6	30,805,426	rs3094127	3.3E-09	1.0E-07	1.3E-06	29,492	Pass				HLA Region		
6	30,834,918	rs3095340	1.8E-04	1.1E-06	4.8E-05	1,421	Pass				HLA Region		
6	30,836,339	rs3094122	1.2E-03	4.8E-04	1.5E-03	18,856					HLA Region		
6	30,855,195	rs10947091	6.4E-03	2.7E-04	2.3E-04	7,324					HLA Region		
6	30,862,519	rs12527415	1.6E-03	1.5E-04	6.7E-05	3,808	Pass				HLA Region		
6	30,866,327	rs11758688	1.7E-03	1.4E-04	5.1E-05	1,677	Pass				HLA Region		
6	30,868,004	rs3131050	1.9E-09	2.1E-09	8.8E-07	3,266	Pass				HLA Region		
6	30,871,270	rs3131060	2.5E-09	4.1E-09	1.2E-06	3,654	Pass				HLA Region		
6	30,874,924	rs4587207	5.0E-02	1.6E-03	9.4E-04	5,433					HLA Region		
6	30,880,357	rs3094123	2.3E-05	1.9E-04	3.7E-03	936	Pass				HLA Region		
6	30,881,293	rs4713376	3.3E-04	6.7E-02	2.5E-02	16,294					HLA Region		
6	30,897,587	rs12195469	5.2E-04	8.0E-02	2.4E-02	13,646					HLA Region		
6	30,911,233	rs3130649	8.5E-04	4.9E-05	4.8E-05	14,612	Pass				HLA Region		
6	30,925,845	rs3095350	2.1E-03	7.3E-05	3.9E-05	77,814	Pass				HLA Region		
6	31,003,659	rs4711247	4.8E-03	1.3E-03	6.9E-04	6,215		DPCR1	31,027,538	31,029,977	HLA Region	DPCR1	
6	31,009,874	rs3131785	1.8E-04	3.1E-03	3.8E-02	2,053		DPCR1	31,027,538	31,029,977	HLA Region	DPCR1	
6	31,011,927	rs3131784	1.8E-04	3.7E-03	4.9E-02	36,439		DPCR1	31,027,538	31,029,977	HLA Region	DPCR1	
6	31,048,366	rs2530710	1.9E-06	1.5E-05	8.6E-05	32,202	Pass	DPCR1	31,027,538	31,029,977	HLA Region	DPCR1	
6	31,080,568	rs1634717	2.8E-04	1.3E-05	4.9E-04	276	Pass				HLA Region		
6	31,080,844	rs1634718	1.1E-06	5.4E-08	3.9E-05	2,784	Pass				HLA Region		
6	31,083,628	rs1632854	2.1E-04	1.2E-05	4.2E-04	37,892	Pass				HLA Region		
6	31,121,520	rs2517538	3.4E-04	1.2E-05	3.1E-03	4,907	Pass				HLA Region		
6	31,126,427	rs2523865	2.1E-09	3.3E-10	5.1E-07	7,265	Pass				HLA Region		
6	31,133,692	rs2517524	4.6E-03	1.9E-02	2.3E-04	32,627					HLA Region		
6	31,166,319	rs3130544	6.2E-17	1.7E-14	5.4E-10	22,131	Pass				HLA Region		
6	31,188,450	rs1265052	2.6E-06	6.5E-07	6.1E-05	1,367	Pass				HLA Region		
6	31,189,817	rs3130975	9.3E-08	2.0E-07	2.2E-05	1,975	Pass				HLA Region		
6	31,191,792	rs3130981	5.8E-06	1.8E-05	1.6E-04	3,320	Pass				HLA Region		
6	31,195,112	rs3095324	1.1E-07	1.2E-07	1.8E-05	2,498	Pass				HLA Region		
6	31,197,610	rs3095314	2.9E-06	2.9E-06	1.2E-04	7,552	Pass				HLA Region		
6	31,205,162	rs3130558	5.3E-06	6.5E-06	7.3E-05	4,491	Pass				HLA Region		
6	31,209,653	rs3130564	4.6E-12	1.0E-10	1.4E-08	28,404	Pass				HLA Region		

6	31,238,057	rs2073723	5.5E-06	1.2E-07	3.3E-04	3,431	Pass						HLA Region	
6	31,241,488	rs2106074	7.0E-08	3.0E-08	4.9E-06	2,197	Pass						HLA Region	
6	31,243,685	rs9263804	6.1E-06	1.5E-07	2.7E-04	747	Pass						HLA Region	
6	31,244,432	rs3130501	4.4E-06	9.7E-08	2.6E-04	261	Pass						HLA Region	
6	31,244,693	rs3132524	6.0E-06	1.4E-07	1.5E-04	6,938	Pass						HLA Region	
6	31,251,631	rs887465	8.6E-05	1.7E-05	1.9E-03	12,133	Pass						HLA Region	
6	31,263,764	rs1265181	3.1E-06	4.6E-06	7.4E-04	5,425	Pass						HLA Region	
6	31,269,189	rs3095238	2.5E-04	2.8E-04	1.9E-03	35,461							HLA Region	
6	31,304,650	rs3130944	9.1E-05	2.0E-05	2.2E-03	8,946	Pass	HLA-C	31,344,508	31,347,834	Strong	Majorcan, Spanish	HLA Region	HLA-C
6	31,313,596	rs3130713	4.9E-05	1.7E-05	3.1E-03	1,303	Pass	HLA-C	31,344,508	31,347,834	Strong	Majorcan, Spanish	HLA Region	HLA-C
6	31,314,899	rs2394894	9.3E-06	4.0E-06	3.6E-04	59	Pass	HLA-C	31,344,508	31,347,834	Strong	Majorcan, Spanish	HLA Region	HLA-C
6	31,314,958	rs2394895	3.6E-05	9.0E-06	4.7E-04	26,531	Pass	HLA-C	31,344,508	31,347,834	Strong	Majorcan, Spanish	HLA Region	HLA-C
6	31,341,489	rs2248880	1.3E-04	1.1E-04	2.9E-03	9,660		HLA-C	31,344,508	31,347,834	Strong	Majorcan, Spanish	HLA Region	HLA-C
6	31,351,149	rs3132486	2.2E-16	4.2E-12	1.8E-10	2,364	Pass	HLA-C	31,344,508	31,347,834	Strong	Majorcan, Spanish	HLA Region	HLA-C
6	31,353,513	rs10456057	8.8E-03	3.1E-03	6.3E-04	202		HLA-C	31,344,508	31,347,834	Strong	Majorcan, Spanish	HLA Region	HLA-C
6	31,353,715	rs6906846	3.8E-09	1.2E-07	2.7E-06	85	Pass	HLA-C	31,344,508	31,347,834	Strong	Majorcan, Spanish	HLA Region	HLA-C
6	31,353,800	rs2524067	4.8E-16	2.8E-13	2.6E-09	1,246	Pass	HLA-C	31,344,508	31,347,834	Strong	Majorcan, Spanish	HLA Region	HLA-C
6	31,355,046	rs7382297	6.0E-04	3.6E-04	3.0E-03	89,183		HLA-C	31,344,508	31,347,834	Strong	Majorcan, Spanish	HLA Region	HLA-C
6	31,444,229	rs2523535	2.5E-05	2.1E-06	1.7E-04	99	Pass	HLA-B	31,429,628	31,432,914	Mixed	Brazilian, Spanish	HLA Region	HLA-B
6	31,444,328	rs2523534	2.2E-06	1.0E-06	2.9E-04	2,737	Pass	HLA-B	31,429,628	31,432,914	Mixed	Brazilian, Spanish	HLA Region	HLA-B
6	31,447,065	rs2596437	4.8E-06	6.5E-06	6.2E-04	4,518	Pass	HLA-B	31,429,628	31,432,914	Mixed	Brazilian, Spanish	HLA Region	HLA-B
6	31,451,583	rs5025315	2.0E-06	3.0E-06	2.6E-04	258	Pass	HLA-B	31,429,628	31,432,914	Mixed	Brazilian, Spanish	HLA Region	HLA-B
6	31,451,841	rs5022119	8.5E-07	2.1E-06	2.1E-04	411	Pass	HLA-B	31,429,628	31,432,914	Mixed	Brazilian, Spanish	HLA Region	HLA-B
6	31,452,252	rs2523638	4.8E-08	8.0E-06	6.0E-04	21	Pass	HLA-B	31,429,628	31,432,914	Mixed	Brazilian, Spanish	HLA Region	HLA-B
6	31,452,273	rs3997982	3.7E-06	2.5E-06	2.9E-04	3,728	Pass	HLA-B	31,429,628	31,432,914	Mixed	Brazilian, Spanish	HLA Region	HLA-B
6	31,456,001	rs2596571	1.0E-06	5.0E-05	1.4E-03	3,013	Pass	HLA-B	31,429,628	31,432,914	Mixed	Brazilian, Spanish	HLA Region	HLA-B
6	31,459,014	rs2523485	1.5E-07	6.4E-06	4.2E-04	1,845	Pass	HLA-B	31,429,628	31,432,914	Mixed	Brazilian, Spanish	HLA Region	HLA-B
6	31,460,859	rs9266774	5.1E-04	3.1E-02	1.6E-02	7,215		HLA-B	31,429,628	31,432,914	Mixed	Brazilian, Spanish	HLA Region	HLA-B
6	31,468,074	rs2596517	3.7E-06	3.3E-06	4.1E-04	40,840	Pass	HLA-B	31,429,628	31,432,914	Mixed	Brazilian, Spanish	HLA Region	HLA-B
6	31,508,914	rs2394999	3.7E-02	1.8E-02	4.6E-04	12,026		MICA	31,479,350	31,491,069	Strong	Eastern India, Basques, Korean, Swedish Caucasians, Czec	HLA Region	MICA
6	31,520,940	rs2596464	8.5E-08	1.7E-06	2.4E-06	12,538	Pass	MICA	31,479,350	31,491,069	Strong	Eastern India, Basques, Korean, Swedish Caucasians, Czec	HLA Region	MICA
6	31,533,478	rs3131621	1.0E-06	2.6E-06	8.3E-06	11,082	Pass	MICA	31,479,350	31,491,069	Strong	Eastern India, Basques, Korean, Swedish Caucasians, Czec	HLA Region	MICA
6	31,544,560	rs2844507	7.1E-06	9.1E-05	2.9E-03	10,129	Pass	MICB	31,573,945	31,586,880			HLA Region	MICB
6	31,554,689	rs2248459	7.5E-15	4.1E-11	2.0E-09	86	Pass	MICB	31,573,945	31,586,880			HLA Region	MICB
6	31,554,775	rs2248462	1.0E-04	4.7E-04	6.4E-05	1,737	Pass	MICB	31,573,945	31,586,880			HLA Region	MICB
6	31,556,512	rs2248617	1.8E-14	3.4E-11	1.3E-09	329	Pass	MICB	31,573,945	31,586,880			HLA Region	MICB
6	31,556,841	rs3749946	2.5E-04	2.3E-02	1.1E-03	114		MICB	31,573,945	31,586,880			HLA Region	MICB
6	31,556,955	rs3099844	2.6E-18	6.9E-15	2.0E-10	105	Pass	MICB	31,573,945	31,586,880			HLA Region	MICB
6	31,557,060	rs2904776	1.3E-04	8.6E-04	2.1E-03	188		MICB	31,573,945	31,586,880			HLA Region	MICB
6	31,557,248	rs2516422	5.1E-06	6.5E-05	1.1E-05	2,207	Pass	MICB	31,573,945	31,586,880			HLA Region	MICB
6	31,559,455	rs2905747	1.6E-04	2.5E-04	1.2E-02	8,667		MICB	31,573,945	31,586,880			HLA Region	MICB
6	31,568,122	rs2395034	1.3E-04	6.9E-04	4.0E-05	125,951	Pass	MICB	31,573,945	31,586,880			HLA Region	MICB
6	31,694,073	rs2736177	3.7E-10	3.2E-08	6.6E-06	16,873	Pass	LTA	31,648,072	31,650,077	Strong	Caucasian, Finnish, Japanese, North Indian Asian	HLA Region	LTA
6	31,710,946	rs1046089	6.1E-15	3.5E-11	1.2E-08	8,810	Pass	LST1	31,661,950	31,664,665	Mixed		HLA Region	LST1
6	31,719,756	rs760293	1.6E-08	3.0E-06	1.5E-04	1,962	Pass	AIF1	31,691,012	31,692,777	Weak		HLA Region	AIF1
6	31,721,718	rs3130048	1.4E-09	4.7E-07	7.9E-04	4,382	Pass	AIF1	31,691,012	31,692,777	Weak		HLA Region	AIF1
6	31,726,100	rs805301	6.6E-15	1.2E-11	1.9E-08	25,401	Pass	AIF1	31,691,012	31,692,777	Weak		HLA Region	AIF1
6	31,751,501	rs9378164	1.4E-02	4.8E-04	2.4E-03	77,511		BAT2	31,696,576	31,713,516	Strong	Spanish	HLA Region	BAT2
6	31,829,012	rs3131379	6.0E-17	1.7E-13	2.4E-09	73,559	Pass	LY6G6C	31,794,404	31,797,489			HLA Region	LY6G6C
6	31,902,571	rs2763979	4.2E-09	1.4E-08	7.3E-07	126,095	Pass	HSPA1B	31,903,503	31,906,010	Weak		HLA Region	HSPA1B
6	32,028,666	rs550513	8.1E-04	6.9E-04	1.2E-02	12,474							HLA Region	
6	32,041,140	rs406936	3.3E-08	1.9E-05	1.8E-04	357	Pass						HLA Region	
6	32,041,497	rs492899	1.4E-05	1.0E-04	6.5E-03	117,025	Pass						HLA Region	
6	32,158,522	rs3130287	2.6E-11	1.5E-08	5.7E-08	9,323	Pass						HLA Region	
6	32,167,845	rs1150753	1.9E-21	7.5E-15	2.5E-10	101,563	Pass						HLA Region	
6	32,269,408	rs204990	2.6E-08	1.8E-07	4.9E-06	422	Pass	NOTCH4	32,270,598	32,299,822			HLA Region	NOTCH4
6	32,269,830	rs204989	1.2E-06	1.2E-06	9.3E-05	3,592	Pass	NOTCH4	32,270,598	32,299,822			HLA Region	NOTCH4
6	32,273,422	rs2071278	1.2E-11	3.1E-09	1.5E-06	14,702	Pass	NOTCH4	32,270,598	32,299,822			HLA Region	NOTCH4
6	32,288,124	rs3131294	9.0E-13	9.0E-09	3.2E-08	8,886	Pass	NOTCH4	32,270,598	32,299,822			HLA Region	NOTCH4
6	32,297,010	rs415929	8.0E-04	9.7E-06	2.1E-06	10,112	Pass	NOTCH4	32,270,598	32,299,822			HLA Region	NOTCH4
6	32,307,122	rs377763	1.7E-18	6.6E-14	8.2E-10	9,180	Pass	NOTCH4	32,270,598	32,299,822			HLA Region	NOTCH4
6	32,316,302	rs424232	4.9E-06	9.4E-05	1.2E-02	4,728	Pass	NOTCH4	32,270,598	32,299,822			HLA Region	NOTCH4
6	32,321,030	rs9267954	1.2E-04	3.6E-03	1.9E-04	20,834		NOTCH4	32,270,598	32,299,822			HLA Region	NOTCH4
6	32,341,864	rs7767325	5.9E-04	2.9E-03	1.7E-02	9,243		NOTCH4	32,270,598	32,299,822			HLA Region	NOTCH4
6	32,351,107	rs1559873	1.2E-04	3.4E-03	8.9E-03	26							HLA Region	
6	32,351,133	rs1559874	1.4E-04	2.2E-03	1.0E-02	37,307							HLA Region	
6	32,388,440	rs9268207	5.0E-05	8.1E-04	8.5E-03	1,048	Pass						HLA Region	
6	32,389,488	rs1018433	1.7E-04	2.5E-03	1.4E-02	24,469							HLA Region	
6	32,413,957	rs3129900	8.7E-15	2.3E-09	5.9E-10	9,675	Pass						HLA Region	
6	32,423,632	rs910050	4.6E-06	3.1E-04	1.6E-03	73	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,423,705	rs910049	8.2E-18	7.7E-12	7.9E-11	10,108	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,433,813	rs2894249	1.9E-18	5.8E-12	8.6E-11	10,292	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,444,105	rs3129932	1.6E-17	1.6E-11	8.6E-11	34	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,444,139	rs3129933	1.3E-06	2.5E-05	3.4E-05	26	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,444,165	rs3129934	3.8E-15	7.2E-10	2.3E-10	5,286	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,449,451	rs9268403	1.9E-08	7.0E-08	3.6E-08	4,216	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,453,667	rs2894254	6.2E-20	2.7E-15	6.9E-11	17,838	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,471,505	rs2076533	2.0E-04	4.3E-04	1.8E-05	289	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,471,794	rs2076530	1.7E-04	3.6E-04	2.7E-05	28	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,471,822	rs9268480	7.3E-09	3.3E-08	1.6E-08	9,341	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,481,163	rs10947260	3.0E-05	2.6E-05	9.9E-06	47	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2

6	32,481,210	rs10947261	5.9E-05	3.4E-05	3.8E-05	1,390	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,482,600	rs3763307	6.3E-09	2.7E-08	1.3E-08	8,788	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,491,388	rs6930933	3.2E-08	2.8E-06	5.2E-07	448	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,491,836	rs2001097	3.8E-08	2.8E-06	3.8E-07	669	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,492,505	rs9268541	7.7E-04	3.9E-02	3.2E-01	572		BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,493,077	rs3135378	3.1E-08	3.3E-06	5.2E-07	300	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,493,377	rs3135377	5.0E-18	2.1E-11	1.7E-10	71	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,493,448	rs3135376	3.9E-08	3.5E-06	5.2E-07	2,282	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,495,730	rs2395161	2.4E-08	3.3E-06	4.8E-07	108	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,495,838	rs2395164	1.4E-09	6.0E-07	1.8E-07	448	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,496,286	rs2395167	1.8E-07	2.2E-06	6.0E-07	266	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,496,552	rs2213580	2.6E-07	2.0E-05	2.7E-06	938	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,497,490	rs9268560	1.2E-05	2.1E-04	3.4E-03	7,103	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,504,593	rs3135342	2.9E-06	6.9E-05	3.5E-06	7,520	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,512,113	rs5000563	2.8E-06	6.7E-05	1.3E-06	3,018	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,515,131	rs3129872	1.8E-06	5.7E-05	1.2E-06	1,374	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,516,505	rs9268645	1.1E-06	1.6E-05	1.6E-04	70	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,516,575	rs3129877	6.5E-06	9.4E-05	1.9E-06	245	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,516,820	rs3135393	6.1E-08	6.0E-06	4.9E-07	3,750	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,520,570	rs1051336	1.8E-07	8.2E-06	1.7E-06	17,051	Pass	BTNL2	32,470,491	32,482,878			HLA Region	BTNL2
6	32,537,621	rs9268853	1.3E-13	6.1E-13	5.9E-11	115	Pass	HLA-DRA	32,515,625	32,520,801	Strong		HLA Region	HLA-DRA
6	32,537,736	rs9268858	5.0E-13	7.9E-13	3.0E-10	1,389	Pass	HLA-DRA	32,515,625	32,520,801	Strong		HLA Region	HLA-DRA
6	32,539,125	rs9268877	9.1E-19	8.9E-16	2.6E-12	113,205	Pass	HLA-DRA	32,515,625	32,520,801	Strong		HLA Region	HLA-DRA
6	32,652,330	rs2027852	1.4E-33	0.0E+00	4.6E-14	29,708	Fail	HLA-DRB5	32,593,129	32,605,984	Mixed	Tunisian, Trinidadian, Thai	HLA Region	HLA-DRB5
6	32,682,038	rs9270986	5.8E-16	2.2E-10	5.2E-11	21,023	Pass	HLA-DRB1	32,654,527	32,665,559	Mixed	African, Japanese, Caucasian, Icelandic, Indian, Majorcan, (	HLA Region	HLA-DRB1
6	32,703,061	rs3129768	1.2E-21	8.7E-14	1.7E-13	7,186	Pass	HLA-DRB1	32,654,527	32,665,559	Mixed	African, Japanese, Caucasian, Icelandic, Indian, Majorcan, (	HLA Region	HLA-DRB1
6	32,710,247	rs9272219	9.3E-05	3.1E-03	5.4E-02	7,158	Pass	HLA-DRB1	32,654,527	32,665,559	Mixed	African, Japanese, Caucasian, Icelandic, Indian, Majorcan, (	HLA Region	HLA-DRB1
6	32,717,405	rs9272723	9.0E-27	0.0E+00	5.8E-14	16,845	Pass	HLA-DQA1	32,713,161	32,719,407	Strong	Russian, African, Taiwanese, European American, Indian, Ice	HLA Region	HLA-DQA1
6	32,734,250	rs9273363	7.1E-89	0.0E+00	0.0E+00	28,368	Pass	HLA-DQA1	32,713,161	32,719,407	Strong	Russian, African, Taiwanese, European American, Indian, Ice	HLA Region	HLA-DQA1
6	32,762,618	rs2856688	2.9E-17	9.8E-15	3.0E-10	3,439	Pass	HLA-DQA1	32,713,161	32,719,407	Strong	Russian, African, Taiwanese, European American, Indian, Ice	HLA Region	HLA-DQA1
6	32,766,057	rs7775228	8.0E-06	5.6E-06	1.5E-05	231	Pass	HLA-DQA1	32,713,161	32,719,407	Strong	Russian, African, Taiwanese, European American, Indian, Ice	HLA Region	HLA-DQA1
6	32,766,288	rs9469220	4.9E-19	4.4E-16	6.6E-13	11,690	Pass	HLA-DQA1	32,713,161	32,719,407	Strong	Russian, African, Taiwanese, European American, Indian, Ice	HLA Region	HLA-DQA1
6	32,777,978	rs2858308	2.2E-05	2.9E-05	2.4E-04	14,029	Pass	HLA-DQB1	32,735,642	32,742,419	Strong	Japanese, Mixageneted population (caucasian, african, ame	HLA Region	HLA-DQB1
6	32,792,007	rs6936863	7.0E-17	3.4E-14	2.4E-11	1,521	Pass	HLA-DQB1	32,735,642	32,742,419	Strong	Japanese, Mixageneted population (caucasian, african, ame	HLA Region	HLA-DQB1
6	32,793,528	rs3916765	5.2E-30	0.0E+00	6.7E-14	3,774	Pass						HLA Region	
6	32,797,302	rs9275765	4.2E-07	2.1E-07	8.5E-06	179	Pass						HLA Region	
6	32,797,481	rs9275772	4.1E-07	2.4E-07	1.1E-05	26	Pass						HLA Region	
6	32,797,507	rs9461799	7.9E-10	8.8E-11	2.2E-09	498	Pass						HLA Region	
6	32,798,005	rs9275793	3.7E-07	1.3E-07	4.7E-06	21,755	Pass						HLA Region	
6	32,819,760	rs2227127	9.0E-11	2.2E-11	1.7E-10	322	Pass						HLA Region	
6	32,820,082	rs9276429	1.2E-10	1.7E-10	1.6E-09	143	Pass						HLA Region	
6	32,820,225	rs9276431	1.2E-10	1.9E-10	1.6E-09	137	Pass						HLA Region	
6	32,820,362	rs9276432	1.2E-10	1.7E-10	1.6E-09	883	Pass						HLA Region	
6	32,821,245	rs2239800	2.3E-06	1.6E-07	6.4E-08	600	Pass						HLA Region	
6	32,821,845	rs9276435	7.8E-04	7.9E-03	1.1E-03	916							HLA Region	
6	32,822,761	rs9276440	3.9E-10	7.5E-10	3.6E-09	15,038	Pass						HLA Region	
6	32,837,799	rs7768538	2.0E-10	1.1E-10	7.9E-10	191	Pass						HLA Region	
6	32,837,990	rs7453920	1.6E-10	1.2E-10	1.4E-09	1,948	Pass						HLA Region	
6	32,839,938	rs6902723	2.9E-08	3.0E-07	6.3E-08	250	Pass	HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,840,188	rs6903130	1.0E-07	7.3E-07	1.3E-07	3,934	Pass	HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,844,122	rs9296044	2.8E-09	1.5E-09	3.7E-10	9,383	Pass	HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,853,505	rs12661352	7.0E-04	2.3E-05	2.2E-02	6,087	Fail	HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,859,592	rs17429127	4.0E-04	3.4E-05	8.5E-03	29,940	Pass	HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,889,532	rs2070121	1.9E-04	7.0E-04	4.9E-04	595		HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,890,127	rs17501267	8.7E-04	1.9E-03	5.2E-03	433		HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,890,560	rs2071474	2.6E-07	3.1E-09	8.0E-07	4,454	Pass	HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,895,014	rs1894407	3.4E-02	1.2E-04	2.9E-03	139		HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,895,153	rs9784858	7.4E-06	5.3E-04	8.1E-03	7,857	Pass	HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,903,010	rs10484565	6.0E-05	5.0E-03	2.4E-02	8,808	Pass	HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,911,818	rs241429	1.2E-02	1.7E-04	1.4E-03	574		HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,912,392	rs241427	1.8E-13	6.6E-11	9.4E-08	62,578	Pass	HLA-DOB	32,888,527	32,892,762			HLA Region	HLA-DOB
6	32,974,970	rs241403	3.1E-09	3.1E-07	6.1E-07	3,065	Pass	TAP1	32,920,964	32,929,726	Weak	Finnish, German, Japanese and Caucasian	HLA Region	TAP1
6	32,978,035	rs3101942	4.8E-12	6.2E-10	3.5E-08	33,843	Pass	TAP1	32,920,964	32,929,726	Weak	Finnish, German, Japanese and Caucasian	HLA Region	TAP1
6	33,011,878	rs151719	8.4E-10	6.9E-07	1.3E-05	13,957	Pass	HLA-DMB	33,010,393	33,016,795	Weak		HLA Region	HLA-DMB
6	33,025,835	rs1050391	9.0E-05	4.0E-04	4.3E-03	123	Pass	HLA-DMB	33,010,393	33,016,795	Weak		HLA Region	HLA-DMB
6	33,025,958	rs11539216	1.6E-04	2.9E-03	1.7E-02	20,219		HLA-DMB	33,010,393	33,016,795	Weak		HLA Region	HLA-DMB
6	33,046,177	rs17840186	4.9E-04	1.5E-02	3.7E-02	27,743		HLA-DMB	33,010,393	33,016,795	Weak		HLA Region	HLA-DMB
6	33,073,920	rs176248	1.8E-06	7.1E-08	4.0E-06	7,801	Pass	HLA-DMA	33,024,373	33,028,831			HLA Region	HLA-DMA
6	33,081,721	rs3129304	1.2E-05	4.5E-06	1.3E-03	135	Pass						HLA Region	
6	33,081,856	rs3129303	1.0E-05	3.0E-06	8.1E-04	4,709	Pass						HLA Region	
6	33,086,565	rs429916	1.8E-08	2.8E-07	3.6E-05	18,990	Pass						HLA Region	
6	33,105,555	rs3135196	2.0E-05	5.7E-06	2.2E-03	27,029	Pass	HLA-DPA1	33,140,772	33,149,356	Mixed		HLA Region	HLA-DPA1
6	33,132,584	rs376877	1.9E-05	2.0E-05	3.6E-03	22,425	Pass	HLA-DPA1	33,140,772	33,149,356	Mixed		HLA Region	HLA-DPA1
6	33,155,009	rs1431403	1.1E-04	2.1E-06	3.1E-04	8,216	Pass	HLA-DPA1	33,140,772	33,149,356	Mixed		HLA Region	HLA-DPA1
6	33,163,225	rs9277542	6.0E-05	3.9E-06	2.2E-04	533	Pass	HLA-DPA1	33,140,772	33,149,356	Mixed		HLA Region	HLA-DPA1
6	33,163,758	rs3128963	3.9E-04	1.9E-05	6.5E-04	166	Pass	HLA-DPA1	33,140,772	33,149,356	Mixed		HLA Region	HLA-DPA1
6	33,163,924	rs3128966	5.9E-03	7.5E-04	1.4E-02	2,928		HLA-DPA1	33,140,772	33,149,356	Mixed		HLA Region	HLA-DPA1
6	33,166,852	rs2179920	2.1E-06	3.5E-07	3.5E-05	11,875	Pass	HLA-DPA1	33,140,772	33,149,356	Mixed		HLA Region	HLA-DPA1
6	33,178,727	rs3128921	2.5E-04	7.0E-06	3.4E-04	573	Pass	HLA-DPA1	33,140,772	33,149,356	Mixed		HLA Region	HLA-DPA1
6	33,179,300	rs3128923	1.2E-04	8.0E-06	1.2E-04	4,313	Pass	HLA-DPA1	33,140,772	33,149,356	Mixed		HLA Region	HLA-DPA1
6	33,183,613	rs3117230	9.3E-07	1.8E-07	1.3E-05	31	Pass	HLA-DPA1	33,140,772	33,149,356	Mixed		HLA Region	HLA-DPA1

6	33,183,644	rs3128930	1.5E-06	2.7E-07	1.6E-05	424	Pass	HLA-DPA1	33,140,772	33,149,356	Mixed				HLA Region	HLA-DPA1
6	33,184,068	rs872956	1.2E-06	5.1E-07	3.3E-06	78,566	Pass	HLA-DPA1	33,140,772	33,149,356	Mixed				HLA Region	HLA-DPA1
6	33,262,634	rs2855459	2.1E-02	5.1E-04	1.2E-02	196,595									HLA Region	
6	33,459,229	rs2747476	4.1E-03	6.6E-04	1.8E-01	186,551										
6	33,645,780	rs17627049	1.7E-02	2.8E-04	1.2E-02	4,721							4,721			
6	33,650,501	rs9296095	2.2E-04	3.1E-03	1.0E-02	32,486		ITPR3	33,697,322	33,772,317	Strong		32,486		ITPR3	ITPR3
6	33,682,987	rs444697	2.5E-04	3.9E-05	5.1E-04	2,373,994	Pass	ITPR3	33,697,322	33,772,317	Strong	ITPR3			ITPR3	ITPR3
6	36,056,981	rs17656051	2.1E-04	1.1E-03	8.9E-04	4,466,734		MAPK14	36,103,551	36,186,513						MAPK14
6	40,523,715	rs10947890	1.5E-02	1.2E-02	4.8E-04	5,600							5,600			LRFN2
6	40,529,315	rs3943783	4.3E-03	6.3E-03	4.5E-04	3,343							3,343			LRFN2
6	40,532,658	rs9471347	9.2E-03	8.2E-03	3.4E-04	2,155							2,155			LRFN2
6	40,534,813	rs1433728	1.7E-03	9.0E-03	2.3E-04	18							18			LRFN2
6	40,534,831	rs2436731	1.2E-02	8.7E-03	4.4E-04	728							728			LRFN2
6	40,535,559	rs987065	8.5E-03	1.9E-02	6.2E-04	3,423							3,423			LRFN2
6	40,538,982	rs1433725	1.3E-02	1.1E-02	4.0E-04	304							304			LRFN2
6	40,539,286	rs10214422	1.7E-02	1.2E-02	7.3E-04	1,813,322										LRFN2
6	42,352,608	rs4714586	6.6E-04	1.5E-01	1.3E-01	688,085										
6	43,040,693	rs2274517	2.6E-04	1.3E-03	1.2E-04	1,727,087										
6	44,767,780	rs489757	1.5E-12	6.8E-07	6.2E-07	1,119,930	Fail					Intergenic				Intergenic
6	45,887,710	rs9395123	3.0E-04	7.7E-03	8.9E-03	931,273										
6	46,818,983	rs12211537	2.7E-04	1.5E-03	6.9E-05	1,814,666	Pass					PLA2G7				PLA2G7
6	48,633,649	rs6910476	8.9E-04	1.3E-02	2.4E-03	153							153			Intergenic
6	48,633,802	rs6458618	4.8E-04	3.2E-03	1.1E-03	28,828							28,828			Intergenic
6	48,662,630	rs9369793	9.7E-04	9.4E-03	2.9E-03	743,379										Intergenic
6	49,406,009	rs542538	1.0E-02	7.8E-04	7.1E-03	2,914,138										
6	52,320,147	rs537288	1.3E-04	3.1E-03	7.8E-03	1,617										
6	52,321,764	rs4715297	3.8E-04	7.6E-03	1.1E-02	89,486										
6	52,411,250	rs3804506	2.1E-04	1.1E-03	1.1E-03	486,537										
6	52,897,787	rs651733	7.5E-22	1.0E-12	5.9E-13	19,328,808	Fail					GSTA3				GSTA3
6	72,226,595	rs9351813	1.0E-02	8.3E-04	1.0E-02	31,884										
6	72,258,479	rs9351816	2.7E-03	3.0E-04	2.9E-03	3,550,686										
6	75,809,165	rs4569931	4.5E-04	3.5E-03	5.3E-02	125,740										
6	75,934,905	rs17783615	8.9E-04	6.6E-04	7.5E-03	45,420										
6	75,980,325	rs1323070	7.4E-04	1.8E-01	1.2E-01	115,289										
6	76,095,614	rs16886475	6.0E-05	2.5E-03	3.5E-04	417,154	Pass					FILIP1				FILIP1
6	76,512,768	rs276699	3.6E-02	3.3E-02	7.2E-04	713,739										
6	77,226,507	rs1490208	1.3E-01	2.2E-01	7.4E-04	7,278										
6	77,233,785	rs7740045	1.3E-01	2.3E-01	8.5E-04	3,262,617										
6	80,496,402	rs9361552	4.8E-04	5.5E-03	1.6E-01	1,792										
6	80,498,194	rs9443691	5.5E-04	6.1E-03	1.6E-01	5,244,595										
6	85,742,789	rs1322872	4.8E-03	6.3E-05	2.5E-04	5,016,370	Pass					Intergenic				Intergenic
6	90,759,159	rs9353711	2.3E-03	2.1E-02	4.3E-04	7,091,598										
6	97,850,757	rs17806588	1.1E-04	4.5E-04	1.9E-03	4,465,017										
6	102,315,774	rs9390779	4.0E-03	5.8E-04	1.1E-01	1,791,613										
6	104,107,387	rs2506743	5.6E-04	7.1E-04	1.1E-02	12,353,897										
6	116,461,284	rs11153594	2.2E-03	8.5E-04	7.9E-03	1,196,003										
6	117,657,287	rs783199	2.1E-02	8.0E-04	4.7E-02	476,145										
6	118,133,432	rs2498587	2.1E-04	3.1E-03	5.7E-03	1,023,784										
6	119,157,216	rs1688618	7.9E-03	1.3E-04	5.9E-02	10,627							10,627			Intergenic
6	119,167,843	rs947138	1.7E-02	2.5E-04	4.3E-02	4,139							4,139			Intergenic
6	119,171,982	rs1417613	1.0E-02	1.1E-04	1.3E-02	9,750,589										Intergenic
6	128,922,571	rs7738739	1.1E-03	8.4E-04	1.8E-02	3,933,122										
6	132,855,693	rs4895939	3.5E-04	2.6E-03	2.6E-02	1,577,373										
6	134,433,066	rs7738638	1.4E-03	1.3E-03	8.3E-04	7,800										
6	134,440,866	rs6906278	4.5E-04	1.6E-03	4.2E-04	3,473,721										
6	137,914,587	rs9689944	7.2E-05	2.0E-05	5.5E-05	494,666	Fail					LOC391040				LOC391040
6	138,409,253	rs9376318	6.5E-05	7.7E-06	2.8E-05	5,008,566	Pass					PERP				PERP
6	143,417,819	rs4896618	9.9E-04	3.0E-03	1.8E-04	3,937,016										
6	147,354,835	rs9322085	2.7E-04	2.1E-04	4.1E-04	38,160							38,160			LOC729178
6	147,392,995	rs1529017	2.6E-03	9.3E-04	3.3E-04	19,763							19,763			LOC729178
6	147,412,758	rs6919814	3.9E-03	2.3E-04	1.8E-03	33							33			LOC729178
6	147,412,791	rs6920096	4.3E-03	2.0E-04	1.5E-03	7,261							7,261			LOC729178
6	147,420,052	rs1881660	7.4E-03	4.9E-04	7.1E-03	18,386							18,386			LOC729178
6	147,438,438	rs6930740	8.8E-03	6.1E-04	5.4E-03	3,382,452										LOC729178
6	150,820,890	rs4870524	4.2E-02	3.4E-04	5.1E-02	351,884										
6	151,172,774	rs782661	2.8E-02	2.3E-02	6.3E-04	109,995										
6	151,282,769	rs9371203	4.4E-05	3.9E-05	3.8E-04	4,900,120	Fail					MTHFD1L				MTHFD1L
6	156,182,889	rs9371392	2.7E-02	1.3E-02	8.5E-04	6,204										
6	156,189,093	rs9384368	1.7E-01	6.4E-02	1.6E-04	60,588										
6	156,249,681	rs4495282	1.1E-03	9.0E-04	6.4E-04	4,248,580										
6	160,498,261	rs9295125	4.2E-02	2.3E-01	8.5E-04	881,394										
6	161,379,655	rs6455701	1.8E-03	1.9E-02	6.6E-04	90,541										
6	161,470,196	rs4709495	1.1E-02	5.6E-02	9.2E-04	201										
6	161,470,397	rs4709496	4.9E-02	2.1E-01	9.1E-04	368,807										
6	161,839,204	rs12191995	7.8E-04	4.0E-03	2.0E-02	36,791							36,791			PARK2
6	161,875,995	rs12195005	3.6E-04	5.0E-03	1.4E-02	4,829							4,829			PARK2
6	161,880,824	rs17651062	9.1E-05	6.5E-04	6.6E-03	5,759,491	Pass					PARK2				PARK2
6	167,640,315	rs4709162	8.1E-04	1.4E-03	7.8E-02	105,546										
6	167,745,861	rs4075645	1.4E-04	6.3E-03	5.6E-02											
7	7,332,182	rs1294610	6.9E-05	5.9E-04	5.6E-05	1,087,287	Poor					COL28A1				COL28A1
7	8,419,469	rs1007095	2.5E-03	3.3E-04	3.3E-02	1,188,983										

7	9,608,452	rs10253648	2.7E-03	1.2E-04	1.1E-02	185					
7	9,608,637	rs10267111	3.4E-03	1.3E-04	1.2E-02	1,686,889					
7	11,295,526	rs2189555	3.9E-04	1.4E-03	2.1E-02	645,450					
7	11,940,976	rs4291160	1.3E-02	7.0E-04	2.1E-03	3,140,475					
7	15,081,451	rs6971261	9.0E-04	2.5E-04	6.0E-03	14,472			14,472		tcag.1136
7	15,095,923	rs9638721	5.5E-04	3.9E-03	6.1E-03	3,729			3,729		tcag.1136
7	15,099,652	rs12699683	3.8E-04	1.0E-03	4.0E-03	344,538					tcag.1136
7	15,444,190	rs12699714	7.4E-04	1.1E-02	7.6E-02	4,077,453					
7	19,521,643	rs2024347	3.5E-05	2.8E-05	2.1E-04	773	Pass	Intergenic	773		Intergenic
7	19,522,416	rs2192486	1.8E-04	1.1E-04	8.3E-04	17,872			17,872		Intergenic
7	19,540,288	rs985307	7.0E-04	5.5E-05	1.3E-04	303	Pass	Intergenic	303		Intergenic
7	19,540,591	rs985306	6.0E-04	3.0E-05	1.4E-04	611	Pass	Intergenic	611		Intergenic
7	19,541,202	rs2192481	1.4E-04	4.6E-05	4.9E-04	828	Pass	Intergenic	828		Intergenic
7	19,542,030	rs985882	1.2E-03	1.3E-04	2.5E-04	50			50		Intergenic
7	19,542,080	rs985881	6.9E-04	4.8E-05	3.9E-05	18,609	Pass	Intergenic	18,609		Intergenic
7	19,560,689	rs2192477	3.6E-05	3.0E-05	2.8E-04	261	Pass	Intergenic	261		Intergenic
7	19,560,950	rs4721790	5.8E-04	7.7E-04	2.1E-04	699			699		Intergenic
7	19,561,649	CG07003679	5.4E-04	1.8E-04	1.1E-04	16,338			16,338		Intergenic
7	19,577,987	rs10248814	1.0E-03	2.9E-04	1.5E-04	4,298			4,298		Intergenic
7	19,582,285	rs7802880	9.0E-05	6.1E-05	8.8E-04	4,891	Pass	Intergenic	4,891		Intergenic
7	19,587,176	rs17451868	8.6E-05	4.6E-05	5.3E-04	1,252	Pass	Intergenic	1,252		Intergenic
7	19,588,428	rs1121732	1.6E-04	4.9E-05	2.5E-04	32,298	Pass	Intergenic	32,298		Intergenic
7	19,620,726	rs4307235	1.3E-04	3.3E-04	8.5E-04	4,582,554					Intergenic
7	24,203,280	rs156293	4.0E-04	1.2E-02	1.9E-03	20,951					
7	24,224,231	rs16217	6.9E-04	3.9E-02	9.9E-04	2,719,486					
7	26,943,717	rs774257	4.5E-04	4.8E-02	5.4E-02	38,487					
7	26,982,204	rs774250	4.8E-04	6.1E-02	7.4E-02	101,069					
7	27,083,273	rs6953314	6.9E-04	9.8E-02	1.1E-01	1,910,664					
7	28,993,937	rs317768	5.2E-02	1.2E-03	5.6E-04	12,311,568					
7	41,305,505	rs37268	1.7E-02	7.6E-04	4.6E-02	11,632,569					
7	52,938,074	rs10256147	1.4E-03	2.0E-03	1.0E-03	185,608					
7	53,123,682	rs6967345	3.9E-04	7.8E-02	7.1E-02	3,229					
7	53,126,911	rs11238249	3.6E-04	8.3E-02	4.9E-02	25,310,608					
7	78,437,519	rs987435	2.2E-03	2.9E-04	6.2E-04	3,369,171					
7	81,806,690	rs11973223	9.7E-05	2.7E-03	3.6E-04	1,356,903	Pass	CACNA2D1			CACNA2D1
7	83,163,593	rs7806361	4.5E-04	9.4E-03	1.0E-02	11,197			11,197		SEMA3E
7	83,174,790	rs505900	4.0E-05	4.1E-03	4.9E-03	19,988	Pass	SEMA3E	19,988		SEMA3E
7	83,194,778	rs17216863	2.5E-04	4.4E-02	3.9E-02	3,060,179					SEMA3E
7	86,254,957	rs10280549	8.6E-02	2.1E-02	2.1E-04	824					
7	86,255,781	rs2299219	9.2E-02	3.1E-02	5.3E-04	2,964,497					
7	89,220,278	rs4728833	2.4E-02	1.2E-01	5.8E-04	416,079					
7	89,636,357	rs2158747	3.5E-02	5.1E-04	3.5E-02	775,193					
7	90,411,550	rs10250625	3.8E-04	4.8E-02	3.4E-03	804,386					
7	91,215,936	rs803271	8.3E-04	1.9E-02	3.3E-02	2,029,706					
7	93,245,642	rs16868514	6.0E-03	2.7E-03	9.0E-04	4,024,085					
7	97,269,727	rs1229556	9.1E-05	8.7E-05	4.6E-04	1,797,644	Fail	ASNS			ASNS
7	99,067,371	rs3735453	1.8E-05	5.4E-04	1.2E-03	5,808,558	Fail	ZNF498			ZNF498
7	104,875,929	rs6977560	3.9E-03	2.0E-04	4.7E-04	271,002					
7	105,146,931	rs4727625	2.9E-02	1.1E-04	2.6E-02	326			326		ATXN7L4
7	105,147,257	rs538249	1.6E-02	4.5E-04	2.4E-02	17,589			17,589		ATXN7L4
7	105,164,846	rs974548	5.3E-04	1.1E-04	7.8E-04	3,930,697					ATXN7L4
7	109,095,543	rs820932	1.5E-02	2.4E-04	1.4E-03	343,388					
7	109,438,931	rs2711371	6.4E-04	6.2E-03	5.2E-03	3,687,480					
7	113,126,411	rs2188422	8.1E-04	1.4E-03	2.1E-02	3,381,855					
7	116,508,266	rs12706141	8.3E-03	7.3E-04	3.8E-02	5,261,879					
7	121,770,145	rs886497	4.4E-03	7.6E-04	5.4E-03	75,817					
7	121,845,962	rs3807874	7.1E-02	9.2E-02	6.2E-04	342,820					
7	122,188,782	rs715205	3.5E-04	2.6E-02	1.3E-03	5,925,828					
7	128,114,610	rs6976879	5.2E-04	9.5E-02	4.5E-03	1,304,707					
7	129,419,317	rs11768743	1.9E-05	2.0E-04	1.0E-03	31,415	Pass	ZC3HC1	31,415		ZC3HC1
7	129,450,732	rs11556924	9.2E-04	2.1E-04	9.2E-04	1,470,592					ZC3HC1
7	130,921,324	rs1643286	9.2E-03	1.5E-04	3.5E-04	2,556,673					
7	133,477,997	rs7782175	3.0E-04	3.3E-02	3.4E-02	12,504					
7	133,490,501	rs997316	3.8E-04	4.2E-02	6.3E-02	7,554,556					
7	141,045,057	rs10261374	4.8E-02	8.0E-04	8.9E-02	138,581					
7	141,183,638	rs7806657	1.5E-04	2.9E-04	1.6E-03	128			128		Olfactory Receptor
7	141,183,766	rs3823668	1.6E-04	4.0E-04	1.2E-03	2,817			2,817		Olfactory Receptor
7	141,186,583	rs7785936	3.0E-04	5.4E-04	9.1E-04	660			660		Olfactory Receptor
7	141,187,243	rs10454325	2.4E-04	4.3E-04	1.1E-03	21,867			21,867		Olfactory Receptor
7	141,209,110	rs1655263	1.0E-04	1.9E-04	1.0E-03	33,015			33,015		Olfactory Receptor
7	141,242,125	rs1726842	2.7E-04	5.4E-04	1.8E-03	11,729			11,729		Olfactory Receptor
7	141,253,854	rs2471744	1.2E-04	3.4E-04	1.1E-03	20,554			20,554		Olfactory Receptor
7	141,274,408	rs1285935	1.2E-04	3.2E-04	1.1E-03	2,634,295					Olfactory Receptor
7	143,908,703	rs7804157	2.9E-03	7.8E-04	9.1E-03	295,516					
7	144,204,219	rs6967341	5.2E-03	6.6E-04	4.8E-03	51,399					
7	144,255,618	rs1404829	7.2E-03	4.3E-04	1.2E-03	643,523					
7	144,899,141	rs6943077	5.7E-03	1.6E-02	4.2E-04	27,876			27,876		Intergenic
7	144,927,017	rs6960519	1.4E-03	4.4E-03	4.5E-04	12,258			12,258		Intergenic
7	144,939,275	rs850498	3.1E-03	6.2E-03	7.4E-04	3,542,871					Intergenic
7	148,482,146	rs917124	3.4E-01	6.5E-02	8.6E-04	3,774,919					
7	152,257,065	rs1860711	1.1E-02	7.7E-04	1.0E-01						

8	2,963,312	rs2045638	8.8E-04	7.6E-03	3.8E-04	1,287,675					
8	4,250,987	rs17336043	1.1E-02	5.7E-04	2.7E-02	768,963					
8	5,019,950	rs7839924	1.8E-02	1.5E-02	1.1E-04	882				882	LOC648237
8	5,020,832	rs12545609	2.2E-02	2.2E-02	2.5E-04	1,383				1,383	LOC648237
8	5,022,215	rs7838077	8.4E-03	3.3E-02	3.2E-04	559,325					LOC648237
8	5,581,540	rs2527120	1.9E-03	2.5E-04	1.4E-02	2,575,857					
8	8,157,397	rs872137	4.4E-02	5.0E-04	3.6E-03	4,523					
8	8,161,920	rs867657	8.6E-02	5.7E-04	2.7E-03	683,082					
8	8,845,002	rs7844524	4.2E-02	6.9E-02	3.1E-04	252,485					
8	9,097,487	rs1053036	5.0E-03	1.7E-03	3.0E-04	25,946					
8	9,123,433	rs4841110	5.1E-03	1.9E-03	1.9E-04	59,059					
8	9,182,492	rs7839174	1.0E-02	8.3E-04	8.0E-03	504,609					
8	9,687,101	rs7829463	1.1E-01	7.7E-04	1.3E-02	1,095,895					
8	10,782,996	rs2898249	2.8E-02	2.0E-03	8.2E-04	5,577,427					
8	16,360,423	rs1464999	2.6E-02	8.6E-04	2.2E-01	10,514				10,514	MRPL49P2
8	16,370,937	rs1526447	3.4E-03	2.5E-04	3.3E-02	11,069				11,069	MRPL49P2
8	16,382,006	rs9325791	1.2E-02	7.5E-04	6.2E-02	2,014				2,014	MRPL49P2
8	16,384,020	rs1526442	4.7E-03	7.1E-04	2.2E-02	506,660					MRPL49P2
8	16,890,680	rs7005198	7.4E-03	3.6E-03	6.3E-04	2,242,874					
8	19,133,554	rs954911	2.9E-04	5.6E-02	1.1E-01	1,051,818					
8	20,185,372	rs4922151	1.4E-03	1.7E-03	1.4E-04	4,326,389					
8	24,511,761	rs11783793	9.3E-04	1.0E-01	1.2E-01	989,599					
8	25,501,360	rs2979734	7.5E-04	3.9E-03	2.7E-04	9,887					
8	25,511,247	rs2979740	5.0E-03	2.1E-02	4.7E-04	2,307,801					
8	27,819,048	rs4515501	4.4E-04	8.4E-03	1.1E-01	1,856,471					
8	29,675,519	rs6558150	1.2E-02	2.7E-03	5.8E-04	2,080,300					
8	31,755,819	rs17660163	4.8E-03	2.8E-04	1.3E-02	10,066					
8	31,765,885	rs2202262	1.1E-02	8.0E-04	2.1E-02	1,666,986					
8	33,432,871	rs2732315	1.7E-03	2.8E-03	2.3E-04	3,466,496					
8	36,899,367	rs1010669	6.3E-02	1.6E-02	3.9E-05	8,484	Pass	FKSG2	8,484		FKSG2
8	36,907,851	rs6986249	4.9E-02	9.4E-03	5.9E-05	80,499	Pass	FKSG2			FKSG2
8	36,988,350	rs3015758	3.9E-02	2.5E-02	8.6E-04	12,586,304					
8	49,574,654	rs6991400	3.2E-04	5.5E-04	2.3E-03	3,906,543					
8	53,481,197	rs1317186	1.9E-02	1.3E-03	8.1E-04	5,749,894					
8	59,231,091	rs11774052	8.5E-05	2.0E-04	1.8E-02	2,865,173	Pass	FAM110B			FAM110B
8	62,096,264	rs16926700	6.3E-04	1.1E-01	5.8E-03	7,365,707					
8	69,461,971	rs16934533	3.9E-09	2.8E-04	4.0E-04	2,901,026	Fail	C8orf34			C8orf34
8	72,362,997	rs4738123	5.0E-03	6.2E-05	8.7E-04	19,878	Pass	EYA1	19,878		EYA1
8	72,382,875	rs920966	1.7E-02	6.9E-05	1.2E-03	640,939	Pass	EYA1			EYA1
8	73,023,814	rs7843557	3.8E-02	4.7E-03	7.2E-04	187,112					
8	73,210,926	rs10504531	9.8E-04	6.1E-03	4.6E-01	1,499,905					
8	74,710,831	rs7817465	4.4E-04	6.1E-03	3.0E-02	704,354					
8	75,415,185	rs10504574	6.1E-04	1.8E-01	1.3E-01	27,072			27,072		GDAP1
8	75,442,257	rs6993852	4.9E-04	1.8E-01	1.9E-01	238			238		GDAP1
8	75,442,495	rs4738452	6.2E-04	1.6E-01	1.7E-01	1,300,834					GDAP1
8	76,743,329	rs9298261	6.0E-03	2.5E-03	7.4E-04	13,420,910					
8	90,164,239	rs1681437	2.1E-03	9.3E-04	1.8E-02	12,068,037					
8	102,232,276	rs17381988	5.4E-03	3.5E-04	8.8E-04	2,183					
8	102,234,459	rs11990866	7.8E-03	7.3E-04	8.8E-04	8,633,328					
8	110,867,787	rs3133919	5.3E-04	1.8E-02	8.8E-03	16,383,139					
8	127,250,926	rs16901179	3.4E-03	9.2E-04	3.0E-03	286			286		Intergenic
8	127,251,212	rs16901182	7.0E-03	8.1E-04	1.9E-03	2,528			2,528		Intergenic
8	127,253,740	CG08023401	2.7E-03	5.8E-04	2.8E-03	243			243		Intergenic
8	127,253,983	CG08023402	1.9E-03	9.5E-04	5.0E-04	1,445,720					Intergenic
8	128,699,703	rs4354269	2.2E-02	7.4E-03	6.8E-04	5,304,771					
8	134,004,474	rs2687810	1.1E-02	7.0E-03	8.8E-04	6,334					
8	134,010,808	rs853296	6.4E-03	3.9E-03	4.2E-04	2,825,613					
8	136,836,421	rs7837809	9.9E-05	1.1E-03	1.9E-02	3,565,837	Pass	Intergenic			Intergenic
8	140,402,258	rs994769	2.7E-03	9.6E-04	1.8E-02						
9	929,738	rs10977650	4.4E-06	6.8E-04	3.0E-03	526,598	Pass	DMRT1			DMRT1
9	1,456,336	rs10961586	1.6E-03	3.9E-04	1.0E-02	460,108					
9	1,916,444	rs7045276	3.3E-04	7.9E-04	1.3E-03	4,769,058					
9	6,685,502	rs1094040	2.7E-04	8.8E-04	5.2E-03	4,350,036					
9	11,035,538	rs10122135	2.3E-02	1.3E-03	9.1E-04	3,340,938					
9	14,376,476	rs662526	1.5E-05	5.5E-04	6.7E-03	6,309	Pass	NFIB	6,309		NFIB
9	14,382,785	rs2208919	2.6E-04	8.4E-03	1.6E-02	497,049					NFIB
9	14,879,834	rs16892	2.5E-02	2.7E-04	1.3E-03	2,704,178					
9	17,584,012	rs7859530	5.5E-04	1.9E-04	1.2E-02	1,573,837					
9	19,157,849	rs2451648	1.0E-03	9.0E-06	4.0E-03	2,660,261	Poor	Intergenic			Intergenic
9	21,818,110	rs3900787	2.8E-02	2.3E-02	6.1E-04	2,897,349					
9	24,715,459	rs7870157	1.1E-04	7.0E-04	5.1E-03	3,339,332					
9	28,054,791	rs2123205	2.4E-02	8.5E-04	9.0E-02	377,552					
9	28,432,343	rs10491888	3.1E-04	2.0E-02	1.6E-02	3,451,801					
9	31,884,144	rs10970584	1.7E-02	3.5E-02	5.8E-04	620					
9	31,884,764	rs7024321	6.8E-03	1.0E-02	3.4E-04	165,013					
9	32,049,777	rs439005	5.1E-03	5.4E-03	2.7E-04	18,398			18,398		Intergenic
9	32,068,175	rs1508318	1.1E-03	8.2E-03	4.1E-04	4,051			4,051		Intergenic
9	32,072,226	rs10970751	2.9E-03	1.2E-02	3.1E-04	11,967			11,967		Intergenic
9	32,084,193	rs10970761	1.4E-03	1.1E-02	5.0E-04	90			90		Intergenic
9	32,084,283	rs10970762	1.7E-03	1.2E-02	2.8E-04	5,596			5,596		Intergenic
9	32,089,879	rs10813733	1.2E-03	7.2E-03	1.6E-04	85,763					Intergenic

9	32,175,642	rs10970830	3.3E-03	4.4E-04	4.9E-05	1,669,496	Pass					Intergenic			Intergenic
9	33,845,138	rs10971738	2.1E-02	1.6E-02	2.5E-04	727,745									
9	34,572,883	rs7032335	8.6E-04	4.5E-03	2.2E-03	36,738									
9	34,609,621	rs10972168	3.6E-04	5.6E-03	6.2E-03	855,770									
9	35,465,391	rs10972486	4.7E-04	1.2E-02	1.4E-02	50,679									
9	35,516,070	rs13298415	6.7E-04	1.1E-03	1.1E-03	416,336									
9	35,932,406	rs7863401	2.2E-03	3.6E-04	1.2E-03	3,569									
9	35,935,975	rs7019283	2.7E-03	4.1E-04	1.6E-03	215,458									
9	36,151,433	rs7851615	6.8E-04	2.6E-02	5.7E-04	2,577,056									
9	38,728,489	rs11794651	8.2E-03	4.4E-04	2.1E-02	628									
9	38,729,117	rs7861857	4.2E-02	7.7E-04	3.8E-02	34,060,528									
9	72,789,645	rs1895063	6.2E-03	4.2E-03	7.3E-04	4,962,801									
9	77,752,446	rs2377425	1.6E-03	2.5E-04	7.2E-04	10,942,416									
9	88,694,862	rs10868500	5.6E-04	6.1E-03	4.3E-02	814,372									
9	89,509,234	rs3118860	4.8E-02	6.0E-02	8.2E-04	7,681							7,681		CTSL
9	89,516,915	rs3128499	6.6E-02	5.6E-02	9.2E-05	11,410	Pass					CTSL	11,410		CTSL
9	89,528,325	rs9410942	1.7E-02	1.3E-02	1.0E-04	4,170							4,170		CTSL
9	89,532,495	rs2274611	1.4E-02	3.4E-02	4.3E-04	11,553							11,553		CTSL
9	89,544,048	rs3128514	2.6E-03	3.3E-03	8.8E-05	8,858	Pass					CTSL	8,858		CTSL
9	89,552,906	rs3118840	9.0E-03	6.7E-02	6.5E-04	2,191,882									CTSL
9	91,744,788	rs7864499	2.0E-04	5.1E-04	2.6E-03	12,260							12,260		IL6RL1
9	91,757,048	rs749194	1.1E-03	9.1E-04	8.1E-03	7,716							7,716		IL6RL1
9	91,764,764	rs10993230	5.7E-04	6.4E-04	6.2E-03	716							716		IL6RL1
9	91,765,480	rs10993257	7.0E-04	6.5E-04	6.2E-03	804							804		IL6RL1
9	91,766,284	rs10821382	8.8E-04	8.6E-04	7.0E-03	495							495		IL6RL1
9	91,766,779	rs10993330	6.7E-04	6.5E-04	5.3E-03	28							28		IL6RL1
9	91,766,807	rs10993331	7.4E-04	9.0E-04	4.6E-03	630							630		IL6RL1
9	91,767,437	rs10761354	6.3E-04	6.3E-04	5.1E-03	1,973							1,973		IL6RL1
9	91,769,410	rs1992544	5.8E-04	6.3E-04	5.2E-03	1,130							1,130		IL6RL1
9	91,770,540	rs10993476	7.7E-04	6.5E-04	6.2E-03	8,526,301									IL6RL1
9	100,296,841	rs10125760	7.4E-04	3.9E-03	2.0E-03	3,388,068									
9	103,684,909	rs9299353	7.7E-03	7.6E-04	1.0E-01	2,604,506									
9	106,289,415	CG09016407	5.9E-02	4.0E-02	3.2E-04	3,811,369									
9	110,100,784	rs10816654	5.8E-04	2.3E-02	3.6E-03	1,163,780									
9	111,264,564	rs3935785	2.0E-02	3.3E-04	5.1E-04	6,468,906									
9	117,733,470	rs6478214	4.0E-04	1.2E-04	2.8E-03	4,266,411									
9	121,999,881	rs10984868	3.4E-03	3.0E-02	4.0E-04	10,749									
9	122,010,630	rs10217145	2.7E-03	6.7E-02	7.5E-04	509,887									
9	122,520,517	rs4617229	1.2E-03	6.6E-04	6.2E-03	5,262,168									
9	127,782,685	rs10987074	3.3E-02	2.8E-04	6.1E-02	5,385									
9	127,788,070	rs1571570	3.4E-02	9.8E-04	7.3E-02	521,597									
9	128,309,667	rs16929088	2.1E-03	5.6E-05	4.0E-02	9,009	Pass					FAM125B	9,009		FAM125B
9	128,318,676	rs4837098	1.4E-03	3.6E-04	3.2E-02	6,806,253									FAM125B
9	135,124,929	rs8176707	2.7E-04	1.5E-02	6.1E-03	2,259,942		ABO	135,120,384	135,140,451	Strong				ABO
9	137,384,871	rs11789038	8.0E-04	2.3E-03	2.4E-01										
10	3,172,026	rs4880595	6.5E-04	7.0E-03	1.2E-02	1,721							1,721		PITRM1
10	3,173,747	rs7094698	8.2E-04	4.0E-03	2.5E-03	5,065							5,065		PITRM1
10	3,178,812	rs7898290	7.3E-04	2.6E-03	8.5E-03	2,715							2,715		PITRM1
10	3,181,527	rs2388557	9.8E-04	9.2E-03	5.1E-03	14,838							14,838		PITRM1
10	3,196,365	rs4620624	7.8E-04	2.2E-02	3.8E-02	2,935,734									PITRM1
10	6,132,099	rs10905669	6.2E-02	1.0E-03	3.9E-02	216,553		IL2RA	6,093,512	6,144,278	Strong			IL2RA	IL2RA
10	6,348,652	rs4750200	2.0E-02	1.6E-02	9.7E-04	2,646,631									
10	8,995,283	rs1822801	5.6E-02	8.6E-04	5.6E-04	1,293,750									
10	10,289,033	rs2759640	2.1E-02	3.2E-03	6.1E-04	836,137									
10	11,125,170	rs1291870	2.5E-03	1.9E-02	3.6E-04	60,089									
10	11,185,259	rs1080845	6.7E-04	7.9E-03	6.1E-02	2,517,758									
10	13,703,017	rs10906425	1.8E-02	7.7E-04	7.1E-03	4,767							4,767		PRPF18
10	13,707,784	rs7089243	9.4E-03	3.1E-04	4.9E-03	11,930							11,930		PRPF18
10	13,719,714	rs1541019	9.2E-03	4.7E-04	6.3E-03	26,659							26,659		PRPF18
10	13,746,373	rs4750396	2.3E-02	7.0E-04	2.5E-02	41							41		PRPF18
10	13,746,414	rs4748050	2.8E-02	8.2E-04	2.0E-02	825,122									PRPF18
10	14,571,536	rs10508474	7.3E-04	1.7E-02	7.8E-03	2,951,657									
10	17,523,193	rs11254600	2.6E-04	1.0E-02	9.8E-02	1,096,550									
10	18,619,743	rs10828424	9.8E-03	2.3E-04	1.5E-02	4,856,772									
10	23,476,515	rs10828412	8.9E-03	6.1E-04	2.8E-03	118,507									
10	23,595,022	rs4074209	3.0E-03	9.9E-04	2.5E-02	5,104,424									
10	28,699,446	rs1249575	8.8E-04	6.1E-03	1.6E-01	1,393,112									
10	30,092,558	rs1013610	8.8E-06	4.4E-05	2.2E-05	57,963	Fail					CKS1BP2			CKS1BP2
10	30,150,521	rs11007760	1.1E-04	1.6E-04	6.7E-03	14,259									
10	30,164,780	rs2150556	1.6E-04	2.8E-04	7.0E-03	577,042									
10	30,741,822	rs6481677	1.7E-02	5.7E-03	5.3E-04	2,665,988		MAP3K8	30,762,872	30,790,767					MAP3K8
10	33,407,810	rs2666275	7.2E-03	9.8E-04	7.0E-03	43,631							43,631		LOC401640
10	33,451,441	rs2070302	1.6E-01	8.5E-03	3.1E-04	12,258							12,258		LOC401640
10	33,463,699	rs2093055	3.7E-03	3.4E-04	3.1E-01	444,681		NRP1	33,506,432	33,663,839					NRP1
10	33,908,380	rs6481861	6.8E-04	2.3E-03	1.3E-03	16,060,147									
10	49,968,527	rs7074818	7.6E-04	2.4E-03	7.5E-03	609,543									
10	50,578,070	rs1258323	5.3E-04	8.5E-03	7.2E-04	3,975,296									
10	54,553,366	rs1006677	2.5E-02	7.8E-04	2.9E-03	2,076							2,076		Intergenic
10	54,555,442	rs10762926	2.8E-02	7.8E-04	2.9E-03	16,283							16,283		Intergenic
10	54,571,725	rs4408237	5.3E-02	9.1E-04	1.5E-03	423							423		Intergenic
10	54,572,148	rs10824900	5.8E-02	9.5E-04	4.2E-03	17,517							17,517		Intergenic

10	54,589,665	rs2384120	3.2E-02	7.4E-04	3.7E-03	16,343,605								Intergenic
10	70,933,270	rs1227969	6.3E-04	4.8E-04	1.0E-02	8,870							8,870	TSPAN15
10	70,942,140	rs10509313	7.7E-04	1.4E-02	3.0E-02	858							858	TSPAN15
10	70,942,998	rs1900463	8.6E-04	1.6E-03	1.0E-03	1,087,067								TSPAN15
10	72,030,065	rs10999427	7.5E-04	6.3E-03	4.2E-03	1,448,408								
10	73,478,473	rs1245574	5.9E-03	1.3E-04	3.5E-02	50,082								
10	73,528,555	rs1668156	3.8E-03	3.2E-04	4.4E-03	73,015								
10	73,601,570	rs11000201	1.2E-04	1.7E-05	1.3E-03	24,088	Pass					ASCC1	24,088	ASCC1
10	73,625,658	rs11000212	6.2E-04	8.7E-05	3.4E-03	6,627,437	Pass					ASCC1		ASCC1
10	80,253,095	rs7098202	4.0E-04	1.7E-02	2.9E-02	299,240								
10	80,552,335	rs704002	1.4E-05	2.1E-05	4.0E-05	139,512	Fail					ZMIZ1		ZMIZ1
10	80,691,847	rs4980054	8.7E-03	1.4E-03	9.8E-05	4,752,489	Pass					ZMIZ1		ZMIZ1
10	85,444,336	rs1931146	1.0E-02	7.9E-04	3.8E-03	4,028,477								
10	89,472,813	rs10887746	3.7E-02	5.8E-02	5.8E-04	1,000,428								
10	90,473,241	rs390436	5.0E-02	7.0E-02	9.8E-04	1,811							1,811	LIPK
10	90,475,052	rs434401	6.1E-02	1.1E-01	7.6E-04	3,392							3,392	LIPK
10	90,478,444	rs427687	4.9E-02	7.3E-02	8.8E-04	9,881,548								LIPK
10	100,359,992	rs526465	6.4E-02	7.4E-04	1.8E-02	4,349,094								
10	104,709,086	rs12413409	3.5E-02	6.3E-02	9.7E-04	1,436,602								
10	106,145,688	rs11192038	1.7E-04	7.4E-03	9.7E-02	476,423								
10	106,622,111	rs2451498	7.3E-03	9.8E-04	1.4E-03	3,540,391								
10	110,162,502	rs1754614	8.7E-04	6.3E-04	2.0E-02	187							187	Intergenic
10	110,162,689	rs1762059	2.2E-04	3.1E-04	7.2E-03	30,949							30,949	Intergenic
10	110,193,638	rs1977486	1.7E-04	2.4E-04	6.6E-03	6,108							6,108	Intergenic
10	110,199,746	rs4528245	6.3E-05	3.8E-04	1.1E-02	53,575	Pass					Intergenic		Intergenic
10	110,253,321	rs1831586	7.6E-04	1.0E-03	3.4E-02	67								
10	110,253,388	rs1332274	2.7E-04	2.8E-04	5.2E-02	7,693,732								
10	117,947,120	rs3824842	4.0E-03	3.6E-02	6.2E-04	4,700,887								
10	122,648,007	rs6585674	7.3E-04	1.2E-02	1.0E-02	1,177,076								
10	123,825,083	rs11200381	1.5E-02	1.5E-03	9.4E-05	279,339	Pass					TACC2		TACC2
10	124,104,422	rs7905027	2.6E-04	3.9E-02	5.1E-03	5,558,210								
10	129,662,632	rs11015984	6.5E-03	2.8E-04	2.1E-02	4,004,188								
10	133,666,820	rs869845	2.2E-03	9.1E-04	7.5E-03	912								
10	133,667,732	rs11596895	1.2E-03	3.9E-04	3.2E-03	125,037								
10	133,792,769	rs1491192	6.9E-03	9.9E-04	2.7E-03	2,917								
10	133,795,686	rs2814184	1.1E-03	6.5E-04	7.1E-03									
11	932,344	rs10751669	1.6E-03	1.5E-03	2.8E-04	3,726							3,726	AP2A2
11	936,070	rs10794350	3.3E-03	2.1E-03	3.4E-04	915							915	AP2A2
11	936,985	rs12278229	1.9E-03	1.6E-03	2.8E-04	3,308							3,308	AP2A2
11	940,293	rs7395402	2.0E-03	1.5E-03	2.8E-04	8,273							8,273	AP2A2
11	948,566	rs10902249	2.3E-03	1.8E-03	3.5E-04	5,414							5,414	AP2A2
11	953,980	rs11246357	2.2E-03	1.8E-03	3.5E-04	291							291	AP2A2
11	954,271	rs7105477	2.6E-03	1.9E-03	3.5E-04	28,420							28,420	AP2A2
11	982,691	rs10902263	7.7E-04	2.3E-03	1.0E-03	871,325								AP2A2
11	1,854,016	rs1806911	2.4E-09	3.1E-04	1.4E-03	927,788	Fail					LSP1		LSP1
11	2,781,804	rs163166	4.4E-05	1.7E-02	3.3E-03	1,582,497	Pass					KCNQ1		KCNQ1
11	4,364,301	rs2269328	3.5E-03	2.9E-04	9.7E-03	1,903,501								
11	6,267,802	rs7104040	6.8E-02	8.0E-04	8.6E-03	11,998,474								
11	18,266,276	rs1520884	7.6E-04	2.9E-02	1.6E-01	47,480							47,480	HPS5
11	18,313,756	rs4150579	8.0E-04	1.2E-02	8.7E-02	8,530							8,530	HPS5
11	18,322,286	rs4150622	8.2E-04	1.8E-02	8.7E-02	17,739							17,739	HPS5
11	18,340,025	rs4150673	8.0E-04	1.4E-02	8.6E-02	1,533,696								HPS5
11	19,873,721	rs11025319	6.1E-05	1.9E-03	2.3E-03	1,033,716	Poor					NAV2		NAV2
11	20,907,437	rs10833412	4.5E-04	8.4E-02	1.9E-02	5,636,272								
11	26,543,709	rs293980	7.9E-03	1.1E-04	5.5E-03	1,590,409								
11	28,134,118	rs1032090	3.5E-01	9.3E-02	8.5E-04	114,013								
11	28,248,131	rs1564745	3.4E-01	8.8E-02	7.2E-04	523								
11	28,248,654	rs1032091	3.5E-01	9.5E-02	8.5E-04	2,037,097								
11	30,285,751	rs10835649	1.3E-02	8.3E-04	3.2E-02	42,440								
11	30,328,191	rs12363432	1.3E-02	5.6E-04	5.4E-02	31,344,208								
11	61,672,399	rs3890102	2.7E-02	5.8E-04	1.2E-01	124,696								
11	61,797,095	rs17709552	1.2E-02	5.1E-04	3.9E-03	6,016,251								
11	67,813,346	rs11606508	2.2E-03	3.1E-03	9.4E-04	1,692,896		LRP5	67,836,711	67,973,317	Weak			LRP5
11	69,506,242	rs10793046	1.7E-02	2.5E-03	2.1E-04	39								
11	69,506,281	rs10793047	9.2E-02	1.4E-02	8.5E-04	9,987,212								
11	79,493,493	rs12802458	4.7E-04	8.7E-03	4.0E-02	3,772,592								
11	83,266,085	rs11600512	6.9E-03	3.5E-03	3.9E-04	12,839							12,839	DLG2
11	83,278,924	rs17146322	7.3E-03	3.9E-03	6.5E-04	27,475							27,475	DLG2
11	83,306,399	rs790363	1.0E-03	8.1E-04	8.9E-04	1,349,257								DLG2
11	84,655,656	rs516620	6.6E-03	9.5E-04	2.6E-02	3,414,916								
11	88,070,572	rs904462	1.0E-02	8.2E-04	3.6E-03	7,516,579								
11	95,587,151	rs525068	8.6E-04	6.7E-02	1.9E-02	11,138,518								
11	106,725,669	rs573449	9.5E-04	6.0E-03	3.1E-02	36,706							36,706	CWF19L2
11	106,762,375	rs660962	8.5E-04	7.0E-03	4.0E-02	29,974							29,974	CWF19L2
11	106,792,349	rs520729	7.1E-04	8.1E-03	3.4E-02	8,429							8,429	CWF19L2
11	106,800,778	rs632910	9.3E-04	3.5E-03	5.1E-02	19,309							19,309	CWF19L2
11	106,820,087	rs652429	6.7E-04	3.6E-03	3.5E-02	434							434	CWF19L2
11	106,820,521	rs654272	5.7E-04	2.0E-03	3.8E-03	8,965							8,965	CWF19L2
11	106,829,486	rs476028	8.2E-04	3.6E-03	3.7E-02	17,264							17,264	CWF19L2
11	106,846,750	rs4753806	4.9E-04	5.8E-03	1.8E-02	1,422							1,422	CWF19L2

11	106,848,172	rs7924419	6.0E-04	1.3E-02	9.6E-02	144							144	CWF19L2
11	106,848,316	rs1943682	5.6E-04	5.7E-03	2.3E-02	1,076,313								CWF19L2
11	107,924,629	rs1940212	7.1E-03	4.8E-04	8.9E-02	13,216							13,216	EXPH5
11	107,937,845	rs2256362	1.3E-03	8.7E-05	1.0E-02	142,277	Pass					EXPH5		EXPH5
11	108,080,122	rs7130825	2.1E-04	3.0E-04	3.8E-03	2,653,186								
11	110,733,308	rs2282637	2.5E-04	2.0E-03	5.8E-03	4,543,710								
11	115,277,018	rs11215719	3.4E-04	4.9E-03	1.4E-02	729,876								
11	116,006,894	rs11216058	1.6E-02	7.5E-04	6.0E-04	700,065								
11	116,706,959	rs609663	2.1E-04	1.4E-02	1.5E-02	3,514,135								
11	120,221,094	rs6589847	7.1E-04	6.4E-03	7.2E-02	4,175,211								
11	124,396,305	rs11219869	9.4E-04	2.3E-01	7.8E-02	31,673							31,673	CCDC15
11	124,427,978	rs11219878	4.4E-04	1.2E-01	1.2E-01	16,466							16,466	CCDC15
11	124,444,444	rs4935903	6.8E-04	1.3E-01	1.3E-01	1,085,299								CCDC15
11	125,529,743	rs4937107	8.5E-03	5.5E-04	6.0E-04	2,568,960								
11	128,098,703	rs947930	1.4E-04	2.9E-02	1.1E-02	680,541								
11	128,779,244	rs4937426	4.8E-04	5.9E-03	3.3E-03	1,034,845								
11	129,814,089	rs10894199	2.5E-03	8.2E-04	2.7E-02	2,802,621								
11	132,616,710	rs4937758	5.2E-03	1.8E-02	5.1E-04	35,899		OPCML	131,790,085	132,907,613				OPCML
11	132,652,609	rs1941217	1.1E-01	7.9E-02	8.8E-04	118,932		OPCML	131,790,085	132,907,613				OPCML
11	132,771,541	rs10894675	4.8E-03	1.0E-03	5.6E-04	205,993		OPCML	131,790,085	132,907,613				OPCML
11	132,977,534	rs11606737	2.2E-03	2.6E-04	1.3E-03	738,205								
11	133,715,739	rs1258852	8.6E-04	2.0E-01	1.0E-01	6,336								
11	133,722,075	rs11223765	2.0E-04	2.4E-03	7.8E-03									
12	1,809,513	rs12301312	1.0E-02	1.4E-02	7.0E-04	375,067								
12	2,184,580	rs12423277	1.8E-02	8.9E-04	8.4E-03	4,591,231								
12	6,775,811	rs1075836	6.0E-02	2.7E-04	5.9E-02	1,878,419		CD4	6,768,912	6,800,237	Strong	Belgian, Danish		CD4
12	8,654,230	rs1561560	3.3E-03	6.8E-03	1.6E-04	2,005,483								
12	10,659,713	rs12306757	1.1E-02	4.3E-04	1.3E-02	6,252								
12	10,665,965	rs2900467	1.6E-02	8.6E-04	5.1E-02	1,434,071								
12	12,100,036	rs2909003	6.3E-08	5.0E-05	4.6E-04	5,035,616	Fail					BCL2L14		BCL2L14
12	17,135,652	rs7974666	4.0E-03	6.4E-04	3.3E-03	302,563								
12	17,438,215	rs10840701	2.5E-02	3.5E-02	1.1E-04	30,967							30,967	LOC390298
12	17,469,182	rs7133850	1.1E-02	3.6E-02	1.0E-03	10,635							10,635	LOC390298
12	17,479,817	rs11043458	2.8E-03	1.5E-02	9.4E-04	7,654,045								LOC390298
12	25,133,862	rs7969931	5.2E-03	4.1E-04	9.9E-03	36,764								
12	25,170,626	rs17387444	6.2E-03	6.9E-04	1.0E-02	975,390								
12	26,146,016	rs12367830	3.6E-04	4.4E-03	1.9E-02	597,395								
12	26,743,411	rs10771294	1.4E-03	6.4E-04	4.5E-04	4,343,128								
12	31,086,539	rs12822390	6.1E-05	2.7E-03	5.7E-03	2,361,952	Poor					TSPAN11		TSPAN11
12	33,448,491	rs7309981	3.3E-02	1.2E-03	8.0E-04	29,657								
12	33,478,148	rs7971226	6.0E-02	4.6E-03	9.8E-04	3,683,630								
12	37,161,778	rs11609286	3.6E-02	1.7E-02	9.3E-04	579							579	Intergenic
12	37,162,357	rs7963716	2.3E-02	1.2E-02	9.0E-04	13,248							13,248	Intergenic
12	37,175,605	rs11168238	1.7E-04	4.9E-03	7.5E-03	49,719							49,719	Intergenic
12	37,225,324	rs11168505	1.8E-02	1.7E-02	7.0E-04	2,381,434								Intergenic
12	39,606,758	rs10784951	2.5E-04	4.1E-02	3.6E-02	3,879								
12	39,610,637	rs11179084	1.6E-04	7.4E-02	2.1E-02	138,463								
12	39,749,100	rs11179687	2.8E-04	4.8E-01	1.4E-01	891,796								
12	40,640,896	rs1234032	1.5E-01	2.8E-04	5.7E-02	53,030								
12	40,693,926	rs1882139	2.5E-01	3.9E-04	3.3E-02	63,541								
12	40,757,467	rs6582378	2.4E-01	3.2E-04	5.6E-02	27,200								
12	40,784,667	rs865420	2.9E-01	7.1E-04	6.0E-02	5,225,758								
12	46,010,425	rs11183856	3.0E-02	7.1E-04	7.6E-02	399,405								
12	46,409,830	rs7484827	9.5E-04	6.2E-04	7.2E-03	4,868,145								
12	51,277,975	rs4103862	5.8E-03	7.0E-04	2.6E-03	1,324,599								
12	52,602,574	rs1443504	3.9E-03	5.4E-04	9.2E-03	152,719								
12	52,755,293	rs4999382	6.3E-04	1.1E-01	2.5E-01	112,038								
12	52,867,331	rs2279402	8.3E-04	1.4E-01	9.5E-02	419								
12	52,867,750	rs2279399	8.9E-04	1.3E-01	1.1E-01	4,431,176								
12	57,298,926	rs885278	1.9E-02	6.8E-04	2.6E-03	4,303								
12	57,303,229	rs2137461	2.3E-02	8.7E-04	3.2E-03	543,477								
12	57,846,706	rs12308619	1.5E-02	5.4E-03	1.2E-04	18,633							18,633	Intergenic
12	57,865,339	rs11172932	4.9E-02	1.7E-02	2.1E-04	1,777							1,777	Intergenic
12	57,867,116	rs11172934	7.4E-02	3.7E-02	2.5E-04	21,053							21,053	Intergenic
12	57,888,169	rs2173646	1.4E-01	4.2E-02	8.3E-04	3,704,833								Intergenic
12	61,593,002	rs772564	4.0E-02	1.0E-02	6.4E-04	517,880								
12	62,110,882	rs11174970	4.5E-03	8.5E-04	1.3E-03	4,367,506								
12	66,478,388	rs7976772	4.0E-04	4.0E-03	3.3E-02	1,515,045								
12	67,993,433	rs1456233	2.6E-03	5.1E-03	4.3E-04	7,432,788								
12	75,426,221	rs12298661	2.4E-04	9.8E-04	6.9E-03	4,425,187								
12	79,851,408	rs7310000	2.9E-04	2.6E-03	2.5E-05	165,438	Pass					ACSS3		ACSS3
12	80,016,846	rs10862237	5.8E-02	3.5E-02	8.0E-04	59,623								ACSS3
12	80,076,469	rs11835646	1.1E-03	9.2E-03	1.0E-04	15,246							15,246	ACSS3
12	80,091,715	rs7964878	2.4E-03	1.2E-02	9.5E-05	289	Pass					ACSS3	289	ACSS3
12	80,092,004	rs7135822	8.5E-03	2.7E-02	5.0E-04	10,653,582								ACSS3
12	90,745,586	rs7304552	9.0E-04	6.9E-02	2.0E-01	53,400								
12	90,798,986	rs10507004	6.8E-03	6.4E-04	2.6E-03	2,108,743								
12	92,907,729	rs669493	3.0E-04	6.8E-03	8.7E-04	770,904								
12	93,678,633	rs6538534	5.6E-04	1.3E-02	8.7E-03	8,369,052								
12	102,047,685	rs1818702	1.5E-03	8.5E-04	2.4E-03	401,685								
12	102,449,370	rs10778260	2.6E-02	1.9E-02	2.7E-04	1,443,454								

12	103,892,824	rs11112269	3.4E-05	3.3E-04	2.3E-03	6,756,274	Pass					KTR18P20			KTR18P20
12	110,649,098	rs634389	1.9E-03	9.6E-04	2.0E-03	25,723									
12	110,674,821	rs6490294	8.0E-03	3.5E-03	8.5E-04	146,924									
12	110,821,745	rs9971746	5.7E-04	4.5E-04	5.3E-04	123,387									
12	110,945,132	rs7114	1.5E-03	9.6E-03	6.9E-04	56,893									
12	111,002,025	rs1980364	9.6E-04	9.4E-03	5.3E-04	190,784		TRAFD1	111,047,764	111,075,795					TRAFD1
12	111,192,809	rs2285809	3.6E-04	1.3E-02	3.6E-04	31,889		C12orf51	111,082,375	111,169,738					C12orf51
12	111,224,698	rs1859246	6.4E-04	2.1E-02	5.0E-04	237,462									
12	111,462,160	rs10850053	6.7E-04	9.2E-03	5.5E-04	292,948		PTPN11	111,340,919	111,432,100					PTPN11
12	111,755,108	rs10850084	1.3E-04	3.4E-03	8.0E-03	2,325									
12	111,757,433	rs10492025	1.4E-04	1.6E-04	6.4E-04	270,778									
12	112,028,211	rs1732771	1.5E-02	9.4E-04	1.2E-01	632,949									
12	112,661,160	rs1114886	9.7E-02	6.7E-05	4.0E-03	2,310,075	Pass					RBM19			RBM19
12	114,971,235	rs3214009	7.3E-04	4.0E-03	7.2E-03	495,870									
12	115,467,105	rs10774871	2.2E-04	1.8E-03	2.7E-03	1,903,600									
12	117,370,705	rs6490178	5.0E-04	1.1E-03	8.4E-02	808,780									
12	118,179,485	rs2727688	5.8E-03	1.6E-03	3.9E-04	1,320,359									
12	119,499,844	rs16950293	1.1E-02	8.1E-04	5.2E-03	7,772,996									
12	127,272,840	rs10847578	9.9E-04	1.1E-02	6.1E-03	1,135									
12	127,273,975	rs9795848	2.2E-03	3.0E-02	5.4E-04	1,811,788									
12	129,085,763	rs628195	1.2E-03	3.6E-04	3.3E-02	314,915									
12	129,400,678	rs3736332	9.7E-04	7.7E-04	5.4E-02	732,908									
12	130,133,586	rs7138155	6.1E-02	6.8E-04	2.1E-02										
13	23,550,171	rs9553161	2.5E-04	1.2E-01	4.6E-02	378									
13	23,550,549	rs912151	5.0E-04	1.6E-01	7.4E-02	1,169,054									
13	24,719,603	rs8262	2.4E-04	7.3E-03	1.5E-03	2,655,037									
13	27,374,640	rs9554196	6.5E-04	1.8E-02	9.5E-03	3,340		PDX1	27,392,177	27,397,410	Mixed		3,340	PDX1	PDX1
13	27,377,980	rs9319399	3.8E-04	1.6E-02	7.5E-03	376		PDX1	27,392,177	27,397,410	Mixed		376	PDX1	PDX1
13	27,378,356	rs9551419	7.4E-04	2.2E-02	9.5E-03	1,706,388		PDX1	27,392,177	27,397,410	Mixed			PDX1	PDX1
13	29,084,744	rs590377	4.6E-04	6.4E-02	4.3E-02	38,942		SLC7A1	28,981,551	29,067,721					SLC7A1
13	29,123,686	rs550984	6.3E-04	7.0E-02	3.7E-02	237,026									
13	29,360,712	rs9551767	8.4E-04	1.7E-01	1.0E-02	1,323									
13	29,362,035	rs9578147	6.8E-04	1.5E-01	7.6E-03	948,093									
13	30,310,128	rs7328020	1.4E-03	4.7E-04	1.7E-02	5,046,703									
13	35,356,831	rs9574777	1.0E-01	3.0E-04	6.3E-02	3,223,201									
13	38,580,032	rs2324129	3.2E-04	2.8E-02	1.1E-03	540,612									
13	39,120,644	CG13004816	1.7E-02	2.9E-03	2.6E-04	4,305,486									
13	43,426,130	rs7331414	3.4E-03	3.4E-02	7.8E-04	2,206							2,206		LOC400126
13	43,428,336	rs7322264	4.3E-03	2.2E-02	9.2E-04	8,016							8,016		LOC400126
13	43,436,352	rs9525895	2.2E-03	1.6E-02	7.8E-04	1,804,878									LOC400126
13	45,241,230	rs9590925	1.6E-02	2.0E-04	5.8E-04	1,515,422									
13	46,756,652	rs7333102	3.0E-03	9.4E-04	1.6E-03	434,390									
13	47,191,042	rs6561402	4.7E-05	1.2E-02	1.3E-02	3,929,181	Pass					LOC730174			LOC730174
13	51,120,223	rs17597531	1.8E-04	3.5E-03	5.4E-02	89,259									
13	51,209,482	rs9535744	2.9E-04	1.0E-02	9.8E-02	14,996							14,996		WDFY2
13	51,224,478	rs12146894	3.7E-04	9.9E-03	6.8E-02	543							543		WDFY2
13	51,225,021	rs9526793	3.8E-04	1.1E-02	9.8E-02	14,161							14,161		WDFY2
13	51,239,182	rs9526795	1.0E-04	6.1E-03	4.3E-02	4,456							4,456		WDFY2
13	51,243,638	rs2296028	4.0E-04	9.1E-03	6.3E-02	1,018,249									WDFY2
13	52,261,887	rs402583	1.1E-01	4.9E-03	7.6E-04	3,572									
13	52,265,459	rs886347	2.0E-02	6.8E-03	3.9E-04	69,940									
13	52,335,399	rs489654	2.2E-03	1.6E-03	2.4E-04	50,760									
13	52,386,159	rs17052520	1.4E-02	3.4E-02	2.8E-04	2,016,241									
13	54,402,400	rs1993777	3.6E-04	3.6E-02	2.8E-02	10,481,498									
13	64,883,898	rs2038825	1.2E-03	6.4E-04	3.6E-02	1,936,375									
13	66,820,273	rs2991432	8.6E-04	7.0E-03	1.3E-02	1,363,320									
13	68,183,593	rs287485	6.8E-04	1.2E-03	3.0E-02	3,519,207									
13	71,702,800	rs7328519	2.4E-04	2.2E-03	1.2E-02	8,263							8,263		Intergenic
13	71,711,063	rs2483837	5.0E-05	3.3E-03	1.0E-02	3,688	Pass					Intergenic	3,688		Intergenic
13	71,714,751	rs340523	7.0E-04	6.7E-03	1.0E-01	170							170		Intergenic
13	71,714,921	rs340524	6.6E-05	2.0E-03	1.4E-02	1,081,145	Pass					Intergenic			Intergenic
13	72,796,066	rs17191385	2.9E-04	3.1E-04	3.2E-04	439,153									
13	73,235,219	rs9318213	1.3E-03	2.0E-03	4.3E-04	1,239,166									
13	74,474,385	rs17337503	1.4E-01	2.4E-02	8.5E-04	5,613,231									
13	80,087,616	rs1772573	1.8E-02	1.8E-04	9.6E-03	283,872									
13	80,371,488	rs6563196	1.1E-04	1.6E-03	5.2E-02	11,877							11,877		Intergenic
13	80,383,365	rs4310745	9.8E-05	1.9E-03	5.3E-02	2,331,766	Pass					Intergenic			Intergenic
13	82,715,131	rs1512771	7.4E-04	8.6E-03	1.0E-01	2,519,561									
13	85,234,692	rs9566093	3.8E-03	2.0E-02	6.3E-04	1,026							1,026		SLITRK6
13	85,235,718	rs9575947	3.2E-03	1.8E-02	5.1E-04	5,870							5,870		SLITRK6
13	85,241,588	rs666540	4.5E-03	2.5E-02	6.1E-04	953							953		SLITRK6
13	85,242,541	rs652258	3.7E-03	2.0E-02	6.1E-04	952							952		SLITRK6
13	85,243,493	rs637400	3.8E-03	1.9E-02	7.6E-04	375							375		SLITRK6
13	85,243,868	rs617544	3.8E-03	2.0E-02	7.6E-04	1,944							1,944		SLITRK6
13	85,245,812	rs588777	3.6E-03	2.2E-02	5.9E-04	17							17		SLITRK6
13	85,245,829	rs606646	3.8E-03	2.5E-02	7.2E-04	33							33		SLITRK6
13	85,245,862	rs606300	3.5E-03	2.0E-02	5.9E-04	381							381		SLITRK6
13	85,246,243	rs694502	4.4E-03	2.7E-02	8.7E-04	1,103							1,103		SLITRK6
13	85,247,346	rs669668	2.1E-03	1.3E-02	5.6E-04	7,320							7,320		SLITRK6
13	85,254,666	rs673100	4.0E-03	1.5E-02	8.4E-04	3,842							3,842		SLITRK6
13	85,258,508	rs598507	4.8E-03	2.6E-02	6.1E-04	46,596							46,596		SLITRK6

13	85,305,104	rs449704	1.3E-03	3.4E-02	7.5E-04	1,952,945					SLITRK6
13	87,258,049	rs1953341	8.1E-04	5.1E-03	1.3E-03	5,627,934					
13	92,885,983	rs7338402	6.8E-03	5.1E-04	3.4E-02	12,122				12,122	GPC6
13	92,898,105	rs9524070	1.1E-02	4.3E-04	3.5E-02	18,471				18,471	GPC6
13	92,916,576	rs895193	2.0E-04	3.2E-05	1.5E-03	51,326	Pass	GPC6			GPC6
13	92,967,902	rs9524111	4.9E-03	2.0E-04	5.8E-03	637,764					GPC6
13	93,605,666	rs1328822	7.4E-04	1.2E-03	4.6E-02	3,616				3,616	GPC6
13	93,609,282	rs9561507	1.9E-04	2.5E-04	9.0E-03	12,798				12,798	GPC6
13	93,622,080	rs16949539	7.9E-04	1.2E-03	1.1E-02	212,699					GPC6
13	93,834,779	rs17196184	6.1E-03	1.1E-03	2.1E-05	32,358	Pass	GPC6		32,358	GPC6
13	93,867,137	rs6650322	3.4E-01	1.1E-01	7.1E-04	2,689,635					GPC6
13	96,556,772	rs2055424	3.4E-04	2.9E-02	2.1E-03	6,473,442					
13	103,030,214	rs12865443	1.2E-01	1.2E-02	9.8E-04	4,656				4,656	Intergenic
13	103,034,870	rs12584285	1.1E-01	1.1E-02	9.8E-04	802				802	Intergenic
13	103,035,672	rs12876017	1.3E-01	1.2E-02	9.8E-04	88,161					Intergenic
13	103,123,833	rs9519133	1.8E-03	4.0E-04	2.0E-04	1,830,289					
13	104,954,122	rs11618600	3.2E-02	1.2E-04	1.8E-02	2,834,883					
13	107,789,005	rs9520845	1.2E-03	2.9E-04	3.3E-02	1,253,397					
13	109,042,402	rs11069790	3.3E-05	2.0E-03	1.5E-03	187,225	Pass	IRS2			IRS2
13	109,229,627	rs4771648	3.3E-03	3.5E-04	1.7E-04	77,719					IRS2
13	109,307,346	rs1041637	1.6E-03	6.0E-06	1.8E-04	572,812	Pass	IRS2			IRS2
13	109,880,158	rs4551896	7.2E-02	2.9E-02	9.4E-04	61,694					
13	109,941,852	rs402661	3.2E-04	1.2E-02	2.1E-03						
14	20,744,204	rs17102539	2.1E-02	9.8E-03	6.6E-04	5,719,075					
14	26,463,279	rs1956073	1.8E-02	8.1E-04	4.0E-02	1,085				1,085	UNGP2
14	26,464,364	rs12588403	6.4E-04	2.4E-03	1.4E-03	10,257				10,257	UNGP2
14	26,474,621	rs1956075	1.0E-02	7.2E-04	4.0E-02	1,220				1,220	UNGP2
14	26,475,841	rs11160092	9.9E-03	3.3E-04	1.8E-02	8,224				8,224	UNGP2
14	26,484,065	rs10431641	6.1E-03	8.1E-04	2.1E-02	3,626				3,626	UNGP2
14	26,487,691	rs12881169	8.6E-05	3.9E-04	2.0E-03	686	Pass	UNGP2		686	UNGP2
14	26,488,377	rs8011593	8.9E-05	2.8E-04	1.4E-03	5,927	Pass	UNGP2		5,927	UNGP2
14	26,494,304	rs2332880	2.1E-04	7.0E-04	2.1E-03	7,007				7,007	UNGP2
14	26,501,311	rs1977165	5.6E-03	6.7E-04	2.1E-02	740				740	UNGP2
14	26,502,051	rs4983208	1.6E-05	1.0E-04	1.1E-03	282,611	Pass	UNGP2			UNGP2
14	26,784,662	rs10873462	5.4E-04	4.9E-03	7.3E-02	94,109					
14	26,878,771	rs7145681	8.8E-04	3.4E-02	7.2E-02	93,495					
14	26,972,266	rs2061262	3.3E-04	4.8E-03	2.5E-02	31,103					
14	27,003,369	rs17112718	3.1E-04	5.4E-04	3.1E-02	4,482,662					
14	31,486,031	rs1278930	8.9E-04	1.7E-04	1.3E-02	626,405					
14	32,112,436	rs8016916	2.0E-04	1.0E-03	5.4E-03	5,357,409					
14	37,469,845	rs8013747	6.8E-04	2.5E-02	1.7E-02	250					
14	37,470,095	rs4597229	6.0E-04	3.2E-02	2.9E-02	16,264,143					
14	53,734,238	rs10143189	3.8E-03	6.4E-02	4.5E-04	1,352,219					
14	55,086,457	rs3742571	1.6E-02	5.4E-03	6.6E-04	2,011,989					
14	57,098,446	rs7156767	2.6E-03	2.3E-04	6.7E-03	2,791,778					
14	59,890,224	rs932606	8.1E-03	1.5E-03	8.1E-04	15,266					
14	59,905,490	rs1254260	4.6E-03	8.4E-04	1.4E-03	7,202,158					
14	67,107,648	rs10138094	3.3E-04	2.8E-03	1.1E-02	4,360					
14	67,112,008	rs4902490	6.3E-04	3.2E-03	6.3E-03	9,227,533					
14	76,339,541	rs2075773	8.7E-02	1.9E-02	7.7E-04	12,763,651					
14	89,103,192	rs10146976	3.5E-04	4.1E-03	4.0E-03	575,768					
14	89,678,960	rs12432204	6.1E-03	5.4E-04	2.0E-02	2,873,091					
14	92,552,051	rs3783904	8.0E-04	6.0E-05	3.0E-04	1,359,343	Fail	ITPK1			ITPK1
14	93,911,394	rs1243167	9.5E-04	7.0E-03	5.2E-03	102,471					
14	94,013,865	rs17825644	3.8E-03	7.1E-04	8.9E-04	6,009					
14	94,019,874	rs8008729	3.2E-03	4.9E-04	6.5E-04	702,004					
14	94,721,878	rs8005908	4.1E-03	2.4E-04	2.2E-03	1,028,449					
14	95,750,327	rs4905462	6.7E-02	1.0E-01	5.7E-04	1,840,758					
14	97,591,085	rs857058	2.2E-02	2.5E-04	2.0E-03	534,991					
14	98,126,076	rs4900406	1.4E-05	2.0E-04	7.5E-04	94,903	Fail	Intergenic			Intergenic
14	98,220,979	rs12432577	5.2E-04	4.0E-04	1.5E-04	1,062,083					
14	99,283,062	rs9324014	9.4E-03	1.0E-02	1.7E-04						
15	22,220,838	rs11161090	3.7E-03	3.2E-02	9.5E-04	2,855,810					
15	25,076,648	rs11636700	2.1E-06	2.4E-04	1.4E-04	11,234	Pass	Intergenic		11,234	Intergenic
15	25,087,882	rs6606891	7.0E-06	4.3E-05	1.3E-03	26,953	Pass	Intergenic		26,953	Intergenic
15	25,114,835	CG15000874	8.6E-04	1.3E-02	6.9E-02	599,102					Intergenic
15	25,713,937	rs3930739	8.5E-04	5.1E-03	1.9E-02	415				415	OCA2
15	25,714,352	rs17674017	6.9E-04	6.6E-03	1.0E-02	7,978				7,978	OCA2
15	25,722,330	rs6497234	5.7E-05	9.8E-04	1.4E-03	2,847	Pass	OCA2		2,847	OCA2
15	25,725,177	rs6497236	1.5E-04	1.6E-03	2.1E-03	26,380				26,380	OCA2
15	25,751,557	rs1391623	7.9E-05	2.8E-04	4.9E-04	474	Pass	OCA2		474	OCA2
15	25,752,031	rs10162623	4.3E-04	8.1E-03	1.2E-02	1,625				1,625	OCA2
15	25,753,656	rs4778192	5.1E-04	9.3E-03	1.4E-02	105				105	OCA2
15	25,753,761	rs4778193	1.3E-04	3.5E-04	6.4E-04	1,444,120					OCA2
15	27,197,881	rs4779798	1.4E-04	5.0E-04	1.2E-04	3,008,156					
15	30,206,037	rs904952	8.0E-04	3.7E-03	2.5E-03	673,989					
15	30,880,026	rs17816387	8.6E-04	2.5E-02	2.5E-02	65,365					
15	30,945,391	rs8040135	6.8E-04	1.2E-03	1.5E-02	8,961				8,961	FMN1
15	30,954,352	rs12591087	2.8E-05	7.3E-04	3.2E-04	14,521	Pass	FMN1		14,521	FMN1
15	30,968,873	rs347937	3.0E-02	2.3E-02	6.0E-04	124,261					FMN1
15	31,093,134	rs11633931	7.2E-03	1.3E-03	2.6E-04	375,795					

15	31,468,929	rs4619367	2.6E-03	1.6E-01	8.2E-04	4,499,789									
15	35,968,718	rs7183305	4.8E-03	2.1E-02	3.6E-04	756,591									
15	36,725,309	rs10438404	6.8E-04	1.0E-01	9.8E-02	1,425,319									
15	38,150,628	rs16970291	3.6E-03	6.5E-04	1.8E-02	2,159,456									
15	40,310,084	rs10518756	4.1E-05	1.5E-04	2.2E-04	5,241,594	Pass				TMEM87A				TMEM87A
15	45,551,678	rs11070587	9.2E-04	6.7E-03	2.0E-04	39,883						39,883			SEMA6D
15	45,591,561	rs12440759	5.9E-03	1.0E-02	7.0E-04	12,105						12,105			SEMA6D
15	45,603,666	rs11070591	3.8E-03	7.4E-03	8.8E-04	16,203						16,203			SEMA6D
15	45,619,869	rs8027222	2.8E-04	3.4E-03	1.3E-04	31,721						31,721			SEMA6D
15	45,651,590	rs12592568	1.6E-04	1.3E-03	7.0E-05	7,983	Pass				SEMA6D	7,983			SEMA6D
15	45,659,573	rs16959700	2.3E-05	1.6E-04	1.2E-05	1,324	Pass				SEMA6D	1,324			SEMA6D
15	45,660,897	rs4774498	2.6E-04	3.4E-03	9.1E-05	18,900	Pass				SEMA6D	18,900			SEMA6D
15	45,679,797	rs12438247	6.5E-05	3.6E-04	3.6E-05	631	Pass				SEMA6D	631			SEMA6D
15	45,680,428	rs16959754	8.7E-06	9.1E-05	6.3E-06	9,872	Pass				SEMA6D	9,872			SEMA6D
15	45,690,300	rs16959825	9.1E-05	7.4E-04	4.4E-05	639	Pass				SEMA6D	639			SEMA6D
15	45,690,939	rs8039597	5.2E-04	3.4E-03	1.4E-04	11,044						11,044			SEMA6D
15	45,701,983	rs1912631	1.6E-02	2.4E-02	8.7E-04	15,730						15,730			SEMA6D
15	45,717,713	rs1435754	5.2E-03	2.2E-03	4.4E-04	2,032						2,032			SEMA6D
15	45,719,745	rs7163692	4.5E-03	1.5E-03	3.6E-04	7,503						7,503			SEMA6D
15	45,727,248	rs10851458	1.7E-02	2.2E-02	9.0E-04	2,509,953									SEMA6D
15	48,237,201	rs8036133	2.9E-03	9.7E-04	1.2E-02	42,831									
15	48,280,032	rs16963406	1.3E-02	6.6E-04	5.4E-02	7,270,724									
15	55,550,756	rs4774943	7.0E-03	4.0E-04	1.7E-02	943,382									
15	56,494,138	rs414846	8.5E-04	1.3E-02	2.2E-02	1,147,926									
15	57,642,064	rs4775177	2.0E-04	5.0E-04	9.7E-03	1,192,345									
15	58,834,409	rs1871858	9.6E-02	5.6E-04	2.1E-02	2,775,945									
15	61,610,354	rs982077	1.2E-02	7.9E-04	2.2E-02	5,262						5,262			USP3
15	61,615,616	rs8025240	1.5E-02	7.7E-04	1.7E-02	3,306						3,306			USP3
15	61,618,922	rs2058913	1.9E-02	7.4E-04	2.2E-02	234,644									USP3
15	61,853,566	rs9972569	1.2E-02	4.5E-04	1.5E-02	922,469									
15	62,776,035	rs16948187	3.8E-04	3.2E-05	1.8E-04	954,912	Fail				OAZ2				OAZ2
15	63,730,947	rs16948916	2.4E-05	1.9E-06	7.4E-04	551,933	Fail				SLC24A1				SLC24A1
15	64,282,880	rs11071870	4.7E-02	1.9E-03	3.9E-04	20,842,888									
15	85,125,768	CG15011200	1.0E-03	2.1E-02	3.0E-01	4,292,529									
15	89,418,297	rs17692973	6.6E-04	4.7E-03	1.6E-02	2,162,716									
15	91,581,013	rs7165051	9.4E-04	3.3E-03	4.7E-04	1,501,897									
15	93,082,910	rs17552225	8.5E-02	8.8E-04	4.4E-01	2,271,528									
15	95,354,438	rs234545	6.1E-03	1.8E-02	5.1E-04	2,486,119									
15	97,840,557	rs1020256	8.4E-04	2.4E-04	6.2E-02										
16	7,104,842	rs13331885	2.4E-04	3.7E-03	2.0E-03	1,418,464									
16	8,523,306	rs11077315	5.2E-03	2.7E-04	2.0E-02	7,968									
16	8,531,274	rs6501156	1.1E-03	8.6E-04	4.6E-03	128,774									
16	8,660,048	rs9923816	2.7E-04	1.9E-03	7.1E-02	4,447,504									
16	13,107,552	CG16003244	2.0E-04	2.7E-02	3.2E-01	2,838									
16	13,110,390	rs17242582	7.3E-04	9.0E-02	4.8E-01	12,718,209									
16	25,828,599	rs4787783	2.5E-03	2.0E-02	9.0E-04	2,294,595									
16	28,123,194	rs7204453	1.6E-02	8.6E-04	4.6E-03	352,985									
16	28,476,179	rs151320	1.0E-02	8.7E-04	7.7E-02	28,206									
16	28,504,385	rs10521145	6.9E-03	6.7E-04	3.0E-02	275,082									
16	28,779,467	rs7201929	3.0E-02	4.4E-04	8.8E-02	877,938									
16	29,657,405	rs7193756	2.1E-04	1.9E-03	3.0E-03	2,890									
16	29,660,295	rs7190132	4.2E-04	2.9E-03	3.7E-03	1,963,807									
16	31,624,102	rs9938037	2.5E-04	1.3E-02	7.3E-03	14,632						14,632			ZNF720
16	31,638,734	rs9928006	5.4E-05	5.8E-03	4.3E-03	114,464	Pass				ZNF720				ZNF720
16	31,753,198	rs8048000	3.2E-04	1.8E-02	3.2E-02	14,702						14,702			VN1R3
16	31,767,900	rs12447780	3.0E-05	9.7E-03	5.2E-03	13,500,470	Pass				VN1R3				VN1R3
16	45,268,370	CG16005854	6.0E-08	4.5E-06	3.1E-05	3,725,654	Fail				VPS35				VPS35
16	48,994,024	rs11648136	1.2E-13	8.0E-08	1.1E-07	287,564	Fail				BRD7				BRD7
16	49,281,588	rs9933594	9.1E-04	1.3E-02	6.3E-03	1,504,658		NOD2	49,288,551	49,324,488			NOD2		NOD2
16	50,786,246	rs2245948	7.3E-03	4.9E-04	5.1E-03	426,121									
16	51,212,367	rs3104805	5.3E-04	5.2E-03	1.6E-01	2,941,423									
16	54,153,790	rs1501999	4.3E-03	1.4E-03	8.4E-04	4,992,168									
16	59,145,958	rs6500133	2.2E-02	9.3E-04	2.3E-02	5,592,996									
16	64,738,954	rs7195549	3.9E-04	8.3E-03	1.2E-01	11,773,565									
16	76,512,519	rs1125693	1.2E-02	5.0E-04	9.9E-03	763,577									
16	77,276,096	rs16948307	8.3E-03	9.6E-04	5.6E-02	497,593									
16	77,773,689	rs8052087	6.2E-04	9.4E-04	1.9E-03	943,371									
16	78,717,060	rs2016178	8.9E-04	1.3E-02	1.6E-02	44,765									
16	78,761,825	rs16944269	9.8E-03	9.2E-02	6.1E-04	2,336,482									
16	81,098,307	rs964745	8.1E-02	1.2E-02	2.2E-04	1,205,169									
16	82,303,476	rs9924248	5.8E-04	1.1E-01	1.3E-02	3,182,433									
16	85,485,909	rs3934493	1.8E-04	3.3E-04	2.7E-04	894						894			Intergenic
16	85,486,803	rs4075602	3.0E-05	6.6E-05	3.1E-04	375,927	Pass				Intergenic				Intergenic
16	85,862,730	rs11649456	6.5E-03	1.3E-04	1.0E-02	1,900									
16	85,864,630	rs13336703	7.6E-03	1.6E-04	6.1E-03	1,290,388									
16	87,155,018	rs885065	1.1E-03	8.0E-04	1.3E-01	10,379									
16	87,165,397	rs11076678	1.8E-02	5.7E-04	8.9E-02	1,166,959									
16	88,332,356	rs1230	6.4E-03	4.6E-04	2.3E-02	26,655						26,655			FANCA
16	88,359,011	rs2159116	1.4E-02	3.9E-05	5.6E-03	4,011	Pass				FANCA	4,011			FANCA
16	88,363,022	rs4785595	1.0E-02	7.2E-04	3.2E-02	3,164						3,164			FANCA
16	88,366,186	rs7190403	4.4E-02	7.3E-05	3.3E-02	621	Pass				FANCA	621			FANCA

16	88,366,807	rs12599180	1.1E-02	6.1E-04	2.8E-02	6,744				6,744	FANCA
16	88,373,551	rs3743859	1.3E-02	6.6E-04	3.1E-02	119,792					FANCA
16	88,493,343	rs10153055	6.8E-02	9.9E-04	1.0E-01	164,128					
16	88,657,471	rs4493039	2.5E-02	2.2E-04	1.6E-01						
17	1,278,556	rs8070640	2.7E-05	9.4E-04	5.7E-05	4,576	Pass	CRK	4,576	CRK	
17	1,283,132	rs8064892	5.6E-05	3.6E-03	1.8E-04	25	Pass	CRK	25	CRK	
17	1,283,157	rs8073032	3.0E-04	1.6E-02	4.2E-04	26,873			26,873	CRK	
17	1,310,030	rs7211083	3.3E-05	2.2E-03	1.0E-04	1,282,591	Pass	CRK		CRK	
17	2,592,621	rs17834974	3.8E-02	1.1E-02	1.0E-04	2,774,805					
17	5,367,426	rs11656183	2.0E-03	1.3E-02	7.7E-04	417,892					
17	5,785,318	rs1183141	8.3E-04	1.3E-03	3.0E-03	1,117,917					
17	6,903,235	rs4796366	5.4E-03	4.6E-03	4.6E-04	2,368,497					
17	9,271,732	rs17808941	1.4E-04	4.2E-02	1.3E-04	5,651					
17	9,277,383	rs9916292	3.6E-04	7.0E-02	1.5E-04	4,563,229					
17	13,840,612	rs9893391	6.6E-03	8.8E-04	1.7E-02	45,915					
17	13,886,527	rs8077837	4.9E-03	5.4E-04	8.5E-03	387,184					
17	14,273,711	rs6502358	1.3E-02	4.8E-04	8.2E-02	422,324					
17	14,696,035	rs11652535	1.2E-04	4.4E-04	3.2E-02	318,971					
17	15,015,006	rs8082487	9.0E-02	9.8E-04	2.8E-02	8,747,009					
17	23,762,015	rs6505081	4.1E-03	7.8E-04	6.2E-05	4,932,612	Pass	SLC46A1		SLC46A1	
17	28,694,627	rs12449721	1.4E-02	9.6E-04	8.1E-02	1,248,264					
17	29,942,891	rs11654085	2.9E-04	9.4E-03	6.8E-03	3,202,256					
17	33,145,147	rs3110648	6.5E-02	9.4E-04	7.9E-02	81,065					
17	33,226,212	rs9900825	2.7E-03	1.0E-03	2.8E-06	949,510	Pass	HNF1B		HNF1B	
17	34,175,722	rs1043515	5.2E-04	1.0E-01	8.0E-04	7,382					
17	34,183,104	rs2338115	3.3E-04	6.2E-02	6.6E-04	4,126,667					
17	38,309,771	rs2593595	1.9E-04	1.0E-02	8.0E-04	1,547,447					
17	39,857,218	rs1869495	7.3E-02	7.5E-04	9.5E-03	4,485,825					
17	44,343,043	rs999475	7.6E-02	5.0E-04	8.1E-02	20,502					
17	44,363,545	rs4793993	3.3E-02	2.2E-04	5.6E-02	1,009,789					
17	45,373,334	rs271661	7.5E-04	4.6E-03	1.2E-01	7,720,314					
17	53,093,648	rs8065120	6.7E-03	1.8E-03	1.1E-04	526,368					
17	53,620,016	rs917606	1.1E-01	1.5E-02	7.0E-04	1,896,265					
17	55,516,281	rs7211063	1.9E-03	4.6E-04	1.6E-03	7,193,754					
17	62,710,035	rs6504495	1.4E-04	6.7E-05	9.9E-03	3,601,756	Pass	PSMD12		PSMD12	
17	66,311,791	rs3859276	5.6E-04	3.5E-03	5.0E-03	5,457					
17	66,317,248	rs16976275	1.5E-04	7.9E-04	1.6E-03	1,220,156					
17	67,537,404	rs2193054	9.1E-04	3.1E-04	1.9E-02						
18	3,732,572	rs3745050	6.5E-03	6.1E-04	1.2E-02	182,768					
18	3,915,340	rs935118	9.0E-04	1.6E-02	1.3E-02	621			621	DLGAP1	
18	3,915,961	rs4076011	6.6E-05	2.1E-03	1.9E-02	157,339	Pass	DLGAP1		DLGAP1	
18	4,073,300	rs9963455	2.8E-04	3.9E-04	5.0E-03	261					
18	4,073,561	rs9955521	3.8E-04	5.5E-04	6.9E-03	2,434,273					
18	6,507,834	rs7504263	9.9E-06	1.6E-04	1.6E-03	924,770	Pass	LOC645387		LOC645387	
18	7,432,604	rs16951745	1.3E-03	4.2E-04	6.1E-03	258			258	PTPRM	
18	7,432,862	rs4798567	4.5E-04	1.2E-02	1.9E-02	5,779			5,779	PTPRM	
18	7,438,641	rs16951794	2.2E-04	4.5E-03	8.1E-03	3,362,835				PTPRM	
18	10,801,476	rs8095854	6.6E-02	1.0E-01	1.6E-04	7,102,886					
18	17,904,362	rs16962762	2.0E-02	7.2E-04	7.2E-03	3,089,416					
18	20,993,778	rs9959583	2.9E-03	4.7E-04	8.7E-03	5,050					
18	20,998,828	rs8085678	6.3E-03	6.3E-04	8.2E-03	590,495					
18	21,589,323	rs1971613	3.1E-03	4.9E-04	2.1E-03	2,813,219					
18	24,402,542	rs16944936	2.7E-02	9.5E-04	1.2E-02	276,907					
18	24,679,449	rs1941195	3.1E-02	3.9E-04	2.1E-03	4,864,390					
18	29,543,839	rs7237757	6.0E-04	6.7E-03	2.5E-02	6,240,719					
18	35,784,558	rs2862294	2.9E-04	5.5E-02	2.6E-03	24,870			24,870	Intergenic	
18	35,809,428	rs17632270	3.2E-04	8.9E-02	5.2E-03	3,808			3,808	Intergenic	
18	35,813,236	rs2568472	1.4E-04	5.3E-02	2.2E-03	14,801			14,801	Intergenic	
18	35,828,037	rs394365	2.8E-04	9.3E-02	5.9E-03	959,659				Intergenic	
18	36,787,696	rs7240599	3.3E-03	2.2E-04	2.9E-03	2,873,517					
18	39,661,213	rs17677611	3.0E-02	8.5E-04	1.0E-02	2,721,636					
18	42,382,849	rs4133965	2.1E-03	4.9E-04	1.2E-02	2,453,249					
18	44,836,098	rs2044550	1.8E-04	2.7E-03	1.0E-01	113,449					
18	44,949,547	rs1943000	9.1E-04	4.9E-03	1.0E-01	55,044					
18	45,004,591	rs12606493	6.9E-04	2.8E-03	8.9E-02	13,220					
18	45,017,811	rs16950465	4.2E-04	2.1E-03	7.5E-02	88,302					
18	45,106,113	rs12606466	3.4E-04	2.9E-02	1.5E-02	44,246					
18	45,150,359	rs12455199	2.7E-04	2.7E-02	1.5E-02	87,659					
18	45,238,018	rs12457899	7.3E-04	5.2E-02	2.0E-02	5,324					
18	45,243,342	rs12456478	6.9E-04	4.3E-02	2.0E-02	206,230					
18	45,449,572	rs12970803	3.3E-04	1.6E-03	3.8E-03	383,581					
18	45,833,153	rs17801909	9.0E-04	5.4E-03	3.7E-03	369					
18	45,833,522	rs17728885	2.5E-04	2.3E-03	7.0E-04	416,034					
18	46,249,556	rs17803514	4.7E-04	1.3E-02	1.0E-02	1,017,362					
18	47,266,918	rs1656747	3.5E-04	2.3E-01	1.6E-01	6,045,862					
18	53,312,780	rs7229018	6.2E-03	3.8E-03	2.8E-04	960,003					
18	54,272,783	rs4940706	6.3E-03	3.1E-04	2.1E-02	1,362,493					
18	55,635,276	rs9951126	1.8E-04	3.3E-02	1.1E-01	5,991,803					
18	61,627,079	rs1994230	3.4E-04	2.1E-03	4.5E-04	4,447,184					
18	66,074,263	rs8088698	5.5E-04	3.9E-02	1.0E-01	2,398,554					
18	68,472,817	rs12962531	2.7E-03	9.5E-04	4.9E-04	18,204					

18	68,491,021	rs9962804	1.4E-01	4.8E-02	6.5E-04	1,860,118					
18	70,351,139	rs12954438	4.2E-04	1.6E-04	1.0E-02	1,584,586					
18	71,935,725	rs9947628	7.6E-04	5.5E-02	3.1E-02	6,057					
18	71,941,782	rs2897285	9.9E-04	6.3E-02	2.2E-02	4,050,652					
18	75,992,434	rs11662921	1.2E-05	2.3E-03	6.4E-03	68	Pass	ADNP2	68	ADNP2	
18	75,992,502	rs11660007	1.0E-05	2.2E-03	8.1E-03	2,172	Pass	ADNP2	2,172	ADNP2	
18	75,994,674	rs1064059	1.4E-05	3.5E-03	9.6E-03	21,121	Pass	ADNP2	21,121	ADNP2	
18	76,015,795	rs7229329	7.9E-05	6.6E-03	2.5E-02	1,102	Pass	ADNP2	1,102	ADNP2	
18	76,016,897	rs3087701	5.5E-05	4.0E-03	2.2E-02	109	Pass	ADNP2	109	ADNP2	
18	76,017,006	rs12201	5.6E-05	5.4E-03	3.0E-02	20,737	Pass	ADNP2	20,737	ADNP2	
18	76,037,743	rs11664816	3.7E-04	7.9E-03	5.3E-02	13,791			13,791	ADNP2	
18	76,051,534	rs8083877	7.3E-04	6.2E-03	1.3E-02						
19	5,872,352	rs274795	4.9E-04	8.7E-03	6.3E-02	1,944,502					
19	7,816,854	rs794442	3.9E-03	1.9E-04	3.6E-05	669,748	Pass	EVI5L		EVI5L	
19	8,486,602	rs2042300	2.3E-03	5.2E-03	2.0E-04	2,708,693					
19	11,195,295	rs4804155	4.6E-03	4.9E-04	2.5E-03	3,269,859					
19	14,465,154	rs8109740	2.4E-04	3.3E-02	3.6E-02	1,408,603					
19	15,873,757	rs2018460	7.9E-05	1.4E-03	4.4E-03	33,893	Pass	CYP4F11	33,893	CYP4F11	
19	15,907,650	rs3810427	3.8E-04	2.0E-02	3.5E-02	17,584,685				CYP4F11	
19	33,492,335	CG19002489	6.4E-04	2.3E-02	6.6E-02	751,115					
19	34,243,450	rs7248438	7.2E-02	1.5E-02	7.2E-04	1,411,994					
19	35,655,444	rs3786803	2.2E-04	2.3E-03	4.7E-02	409,530					
19	36,064,974	rs1466297	1.1E-02	1.7E-02	4.9E-04	2,268,622					
19	38,333,596	rs10423969	4.5E-03	5.5E-04	3.1E-03	2,562,934					
19	40,896,530	rs107068	2.0E-02	1.4E-04	1.9E-03	7,178,524					
19	48,075,054	rs10422560	9.3E-05	1.0E-02	1.6E-03	5,850,164	Poor	PSG1		PSG1	
19	53,925,218	rs281408	3.1E-04	6.8E-05	2.3E-06	16,292	Pass	IZUMO1	16,292	IZUMO1	
19	53,941,510	rs12975781	7.2E-06	2.6E-05	1.8E-07	2,969,260	Pass	IZUMO1		IZUMO1	
19	56,910,770	rs11667974	2.2E-05	3.7E-05	6.8E-05	2,633,546	Fail	HAS1		HAS1	
19	59,544,316	rs4356595	5.6E-04	4.7E-04	6.4E-02						
20	2,391,441	rs2143862	2.4E-03	6.7E-03	2.8E-04	8,010					
20	2,399,451	rs6049288	8.6E-03	3.1E-02	6.7E-04	1,920,995					
20	4,320,446	rs6084726	9.1E-03	3.5E-04	1.2E-01	1,656,871					
20	5,977,317	rs6076920	3.8E-03	3.5E-04	8.4E-04	5,915,708					
20	11,893,025	rs1997812	1.3E-02	3.0E-02	3.3E-04	4,618,247					
20	16,511,272	rs6135804	4.4E-04	1.7E-03	8.3E-04	66,207					
20	16,577,479	rs6111223	7.1E-04	2.9E-02	2.7E-01	2,260,206					
20	18,837,685	rs6136562	1.0E-03	1.3E-02	1.1E-02	64,764					
20	18,902,449	rs6035176	3.8E-04	4.2E-03	2.4E-02	4,287,318					
20	23,189,767	CG20005607	1.4E-01	6.7E-02	8.2E-04	10,180,808					
20	33,370,575	rs6060369	3.1E-01	7.0E-01	6.0E-04	7,047			7,047	UQCC	
20	33,377,622	rs6060373	3.0E-01	6.1E-01	3.9E-04	44,045			44,045	UQCC	
20	33,421,667	rs1570004	2.5E-01	4.8E-01	2.8E-04	13,800			13,800	UQCC	
20	33,435,467	rs4911496	2.4E-01	4.6E-01	1.9E-04	129,768				UQCC	
20	33,565,235	rs2236160	5.8E-02	3.8E-01	7.1E-04	6,392,944					
20	39,958,179	rs2076395	5.0E-04	3.1E-03	1.8E-03	218			218	LOC643172	
20	39,958,397	rs6129968	3.4E-04	2.3E-03	8.0E-04	30,152			30,152	LOC643172	
20	39,988,549	rs4810344	9.5E-04	2.3E-03	4.4E-03	10,590			10,590	LOC643172	
20	39,999,139	rs4810345	7.4E-04	3.2E-03	3.8E-03	5,593,887				LOC643172	
20	45,593,026	rs1212595	1.9E-04	2.3E-02	2.4E-04	1,354,944					
20	46,947,970	rs6019503	7.4E-04	7.3E-02	2.7E-02	32					
20	46,948,002	rs6019504	2.9E-04	2.4E-02	8.6E-03	3,098,768					
20	50,046,770	rs6068038	5.3E-05	3.9E-03	4.5E-03	748,104	Pass	Intergenic		Intergenic	
20	50,794,874	rs241800	5.4E-03	9.7E-04	3.2E-02	2,729			2,729	Intergenic	
20	50,797,603	rs17806723	1.9E-04	4.1E-04	7.6E-03	1,310			1,310	Intergenic	
20	50,798,913	rs17734029	1.1E-04	3.2E-04	7.2E-03	4,281			4,281	Intergenic	
20	50,803,194	rs17806747	4.4E-04	5.9E-04	7.6E-03	980,105				Intergenic	
20	51,783,299	rs6091737	4.2E-03	4.9E-04	9.1E-02	2,396,245					
20	54,179,544	rs6014619	1.9E-01	7.4E-04	1.3E-02	403,215					
20	54,582,759	rs6069788	2.3E-04	2.5E-03	2.0E-03	308,279					
20	54,891,038	rs7261954	7.1E-04	3.4E-02	2.1E-03	659					
20	54,891,697	rs6025234	7.5E-04	2.8E-02	1.5E-03	1,362,133					
20	56,253,830	rs3787158	1.4E-02	1.5E-02	6.2E-04	278					
20	56,254,108	rs595292	9.0E-02	2.5E-02	7.7E-04	720,069					
20	56,974,177	rs235189	6.6E-04	2.1E-02	3.3E-02	1,644,211					
20	58,618,388	rs6027712	6.4E-03	1.4E-02	6.8E-04	228,430					
20	58,846,818	rs6015858	1.0E-02	6.4E-04	9.1E-03	1,079,641					
20	59,926,459	rs6061892	8.6E-04	2.7E-04	5.5E-05	871,132	Poor	CDH4		CDH4	
20	60,797,591	rs6090338	4.4E-04	2.4E-04	8.1E-05		Pass	SLC04A1		SLC04A1	
21	26,266,108	rs2051174	1.2E-01	1.2E-01	7.5E-04	2,497,669					
21	28,763,777	rs2831839	2.7E-03	6.6E-04	6.5E-02	189,124					
21	28,952,901	rs2831951	7.2E-03	6.9E-02	4.4E-04	5,312,903					
21	34,265,804	rs2834310	7.0E-04	8.5E-04	1.5E-01	671,708					
21	34,937,512	rs2834556	4.4E-02	1.1E-01	7.5E-04	1,565,285					
21	36,502,797	rs2156412	1.3E-04	3.1E-02	8.6E-02	2,386,149					
21	38,888,946	rs17194186	4.8E-02	1.9E-03	1.9E-04	1,406			1,406	ERG	
21	38,890,352	rs2836550	6.9E-02	2.7E-03	5.6E-04	52			52	ERG	
21	38,890,404	rs2212596	1.0E-02	6.2E-04	4.6E-03	17,711			17,711	ERG	
21	38,908,115	rs460574	4.2E-02	8.1E-04	2.3E-02	339,956				ERG	
21	39,248,071	rs9636950	6.3E-05	2.2E-03	1.1E-02	15,681	Pass	FLJ45139	15,681	FLJ45139	
21	39,263,752	rs2836792	3.4E-04	8.1E-03	1.5E-02	3,458			3,458	FLJ45139	

21	39,267,210	rs2836795	2.8E-05	2.7E-03	4.3E-03	51	Pass	FLJ45139	51	FLJ45139	
21	39,267,261	rs2836796	3.5E-05	3.2E-03	2.7E-03	93	Pass	FLJ45139	93	FLJ45139	
21	39,267,354	rs2836797	9.9E-05	6.1E-03	4.9E-03	4,177	Pass	FLJ45139	4,177	FLJ45139	
21	39,271,531	rs2836801	8.1E-05	5.3E-03	1.3E-02	99,228	Pass	FLJ45139		FLJ45139	
21	39,370,759	rs17231256	5.2E-04	4.5E-03	5.6E-03	486,752				FLJ45139	
21	39,857,511	rs571227	9.3E-02	3.0E-01	9.9E-04	2,288			2,288	FLJ45139	
21	39,859,799	rs523372	1.7E-01	3.6E-01	8.9E-04	8,195			8,195	FLJ45139	
21	39,867,994	rs1734930	1.6E-01	9.6E-02	4.2E-04	2,638,268				FLJ45139	
21	42,506,262	rs221948	6.9E-04	3.9E-02	1.5E-01	573,947					
21	43,080,209	rs883471	7.8E-04	1.3E-01	1.7E-01	1,917,610					
21	44,997,819	rs690260	3.3E-04	3.7E-03	9.5E-04	1,842,295					
21	46,840,114	rs2839342	3.6E-06	1.9E-05	3.4E-04		Fail	S100B		S100B	
22	19,542,090	rs178067	6.6E-04	1.0E-03	2.4E-05	5,096,146	Fail	SNAP29		SNAP29	
22	24,638,236	rs1476053	2.5E-08	5.3E-06	3.9E-05	3,186,388	Fail	MYO18B		MYO18B	
22	27,824,624	rs2294242	7.7E-03	3.6E-04	7.3E-02	144,466					
22	27,969,090	rs5763087	3.1E-06	1.2E-05	1.4E-03	549,155	Poor	EMID1		EMID1	
22	28,518,245	rs5752962	9.7E-04	2.9E-02	5.9E-03	62,067					
22	28,580,312	rs5752972	3.0E-04	3.9E-02	8.5E-03	63,780					
22	28,644,092	rs713718	1.4E-04	2.6E-02	4.4E-03	79,411					
22	28,723,503	rs5763688	2.5E-04	1.5E-02	9.5E-03	12,480			12,480	MTMR3	
22	28,735,983	rs41158	2.0E-05	1.5E-03	2.0E-03	1,381	Pass	MTMR3	1,381	MTMR3	
22	28,737,364	rs41159	4.2E-04	2.1E-02	1.2E-02	25,039			25,039	MTMR3	
22	28,762,403	rs41176	3.5E-04	2.2E-02	1.6E-03	39,851			39,851	MTMR3	
22	28,802,254	rs1989870	2.4E-04	1.5E-02	2.7E-03	58,837				MTMR3	
22	28,861,091	rs4820830	6.0E-04	3.5E-02	7.5E-03	36,816					
22	28,897,907	rs2412976	4.2E-04	2.6E-02	4.5E-03	5,892,113					
22	34,790,020	rs5995203	5.0E-03	2.7E-04	4.7E-05	1,350,633	Pass	APOL3		APOL3	
22	36,140,653	rs133738	2.8E-04	1.3E-02	5.2E-02	1,463					
22	36,142,116	rs133736	5.6E-04	3.7E-02	6.1E-02	1,687,434					
22	37,829,550	rs139314	2.0E-03	9.7E-03	6.9E-04	8,068,718					
22	45,898,268	rs738662	4.7E-03	2.3E-04	5.8E-02	11,341					
22	45,909,609	rs5767523	3.2E-03	4.4E-04	4.9E-02	3,526,470					
22	49,436,079	rs6009945	2.1E-04	3.0E-03	1.6E-02						
23	11,715,814	rs5935164	8.6E-04	2.1E-02	1.2E-01	4,763,948					
23	16,479,762	rs2428447	3.4E-02	5.6E-02	5.3E-04	641,837					
23	17,121,599	rs7049284	7.5E-04	1.4E-02	5.2E-04	10,344,080					
23	27,465,679	rs221406	8.2E-04	2.0E-03	1.9E-02	5,041,354					
23	32,507,033	rs3928369	9.0E-04	9.6E-03	1.3E-02	1,069,617					
23	33,576,650	rs2382963	9.1E-04	6.0E-03	2.2E-02	47,480,190					
23	81,056,840	rs4131395	2.1E-04	5.3E-04	1.1E-02	833,095					
23	81,889,935	rs4828519	5.2E-04	2.6E-03	9.6E-04	1,668,266					
23	83,558,201	rs4437785	1.0E-05	2.3E-05	8.0E-04	667,407	Pass	HDX		HDX	
23	84,225,608	rs5923256	1.9E-04	3.4E-03	5.8E-03	730,966					
23	84,956,574	rs2206006	1.8E-05	6.6E-03	3.1E-02	409,559	Pass	CHM		CHM	
23	85,366,133	rs5923512	6.8E-04	1.9E-02	1.3E-02	590,578					
23	85,956,711	rs11797054	4.7E-04	9.3E-03	1.3E-03	289,260					
23	86,245,971	rs5923792	1.8E-07	4.8E-06	5.1E-04	394,107	Pass	Intergenic		Intergenic	
23	86,640,078	rs2213671	2.5E-06	2.5E-04	2.8E-02	778,740	Pass	KLHL4		KLHL4	
23	87,418,818	rs5969586	2.7E-04	4.9E-03	3.0E-02	96,500					
23	87,515,318	rs6522039	1.6E-04	8.4E-04	5.8E-03	82,174					
23	87,597,492	rs12842964	2.7E-04	2.0E-03	2.5E-02	469,070					
23	88,066,562	rs682176	3.0E-06	7.2E-04	8.8E-04	73,854	Pass	Intergenic		Intergenic	
23	88,140,416	rs3850155	1.5E-05	5.8E-04	5.5E-03	151,629	Pass	Intergenic		Intergenic	
23	88,292,045	rs5942420	2.9E-05	1.8E-03	1.3E-02	3,026,341	Pass	Intergenic		Intergenic	
23	91,318,386	rs4358964	9.2E-05	2.2E-02	8.7E-02	519,886	Pass	PCDH11X		PCDH11X	
23	91,838,272	rs5942313	9.9E-06	3.4E-03	1.2E-04	469,497	Pass	PCDH11X		PCDH11X	
23	92,307,769	rs785754	5.8E-06	1.7E-06	4.8E-05	166,584	Pass	LOC401602		LOC401602	
23	92,474,353	rs5940144	2.4E-08	1.4E-07	6.5E-05	369,783	Pass	LOC643310		LOC643310	
23	92,844,136	rs2312033	4.3E-04	3.1E-02	6.7E-02	158,907					
23	93,003,043	rs1797020	4.5E-04	1.5E-03	1.2E-01	618,270					
23	93,621,313	rs2782710	2.2E-04	1.3E-03	1.3E-01	607					
23	93,621,920	rs1582247	2.1E-04	1.9E-03	1.1E-01	640,017					
23	94,261,937	rs7061674	6.1E-06	3.3E-03	2.4E-03	59,436	Pass	Intergenic	59,436	Intergenic	
23	94,321,373	rs6619774	9.9E-06	1.7E-06	1.3E-05	16,569	Pass	Intergenic	16,569	Intergenic	
23	94,337,942	rs10855336	9.1E-03	6.2E-04	5.8E-04	22,841			22,841	Intergenic	
23	94,360,783	rs5950249	6.6E-07	2.5E-05	3.3E-06	410,820	Pass	Intergenic		Intergenic	
23	94,771,603	rs2153931	1.1E-04	1.3E-04	2.2E-03	1,175,421					
23	95,947,024	rs233655	3.9E-04	6.2E-03	7.3E-03	1,360,032					
23	97,307,056	rs1458782	4.5E-06	4.3E-03	2.8E-04	53,894	Pass	LOC392505	53,894	LOC392505	
23	97,360,950	rs1458780	5.4E-06	2.7E-03	4.4E-04	2,394,686	Pass	LOC392505		LOC392505	
23	99,755,636	rs736550	3.8E-06	2.8E-03	1.3E-02	1,465,954	Pass	TSPAN6		TSPAN6	
23	101,221,590	rs5944840	7.0E-04	1.9E-02	3.2E-03	1,666,251					
23	102,887,841	rs532860	7.2E-06	1.5E-05	3.6E-04	5,810,170	Pass	PLP1		PLP1	
23	108,698,011	rs942602	3.4E-04	1.5E-02	2.4E-02	461,501					
23	109,159,512	rs2066364	4.3E-06	1.5E-06	9.1E-06	2,708,220	Pass	TMEM164		TMEM164	
23	111,867,732	rs16987072	9.2E-05	2.4E-03	1.1E-03	194,988	Pass	LOC203510		LOC203510	
23	112,062,720	rs5974328	5.5E-05	6.5E-04	1.1E-05	123,665	Pass	Intergenic		Intergenic	
23	112,186,385	rs3125994	8.6E-07	2.5E-05	4.5E-04	461,253	Pass	Intergenic		Intergenic	
23	112,647,638	rs5929302	1.1E-05	9.9E-04	3.6E-05	458,450	Pass	LOC139466		LOC139466	
23	113,106,088	rs6642665	1.2E-04	3.8E-04	4.8E-04	2,603,900					
23	115,709,988	rs5909877	2.6E-05	7.6E-04	1.9E-03	41,418	Pass	LOC653155	41,418	LOC653155	

23	115,751,406	rs17280008	7.3E-04	8.2E-03	7.9E-03	221,651					LOC653155
23	115,973,057	rs6646568	5.9E-06	1.8E-03	2.8E-03	39,012	Pass	Intergenic	39,012		Intergenic
23	116,012,069	rs4462068	1.1E-06	9.1E-03	1.1E-04	631,197	Pass	Intergenic			Intergenic
23	116,643,266	rs2878761	2.0E-04	2.1E-02	1.0E-02	17,805					
23	116,661,071	rs5912059	6.9E-04	1.2E-02	2.4E-03	129,799					
23	116,790,870	rs2215785	1.3E-05	4.2E-03	5.2E-03	538,714	Pass	KLHL13			KLHL13
23	117,329,584	rs1716765	1.8E-04	5.1E-04	6.1E-02	98,400					
23	117,427,984	rs5956974	3.4E-04	1.2E-02	2.5E-03	58,424					
23	117,486,408	rs5956063	7.2E-04	3.2E-03	8.3E-04	130,996					
23	117,617,404	rs5910392	5.8E-05	3.1E-05	1.3E-03	1,743,170	Pass	DOCK11			DOCK11
23	119,360,574	rs5910840	1.8E-06	4.3E-04	2.0E-01	339,656	Pass	ATP1B4			ATP1B4
23	119,700,230	rs4825716	3.4E-07	2.9E-06	9.2E-06	741,224	Pass	C1GALT1C1			C1GALT1C1
23	120,441,454	rs6608252	3.3E-11	2.9E-07	2.4E-04	226,422	Pass	Intergenic			Intergenic
23	120,667,876	rs2507247	1.0E-04	2.9E-05	2.3E-03	623,489	Pass	Intergenic			Intergenic
23	121,291,365	rs6648753	3.1E-05	2.0E-02	5.1E-02	39,452	Pass	Intergenic	39,452		Intergenic
23	121,330,817	rs12387318	6.4E-06	1.9E-02	2.1E-02	709,591	Pass	Intergenic			Intergenic
23	122,040,408	CG23007712	7.1E-04	9.9E-04	5.4E-05	1,513	Pass	RP1-57A13.2	1,513		RP1-57A13.2
23	122,041,921	rs924159	5.2E-04	1.8E-03	2.7E-05	252,298	Pass	RP1-57A13.2			RP1-57A13.2
23	122,294,219	rs10521718	1.3E-04	1.7E-02	3.1E-02	121,033					
23	122,415,252	rs6608087	7.1E-04	2.9E-03	1.5E-02	202,700					
23	122,617,952	rs5958281	7.1E-09	2.5E-04	6.7E-05	660,357	Pass	THOC2			THOC2
23	123,278,309	rs1292210	4.8E-04	6.6E-03	5.7E-02	15,898					
23	123,294,207	rs5911809	5.1E-04	8.1E-03	6.4E-02	54,302					
23	123,348,509	rs5956652	8.0E-04	4.4E-03	1.3E-04	24,249					
23	123,372,758	rs2239477	2.8E-04	1.6E-02	1.7E-03	1,644,003					
23	125,016,761	rs17266022	9.4E-05	1.1E-03	1.4E-03	117,143	Pass	WDR40C			WDR40C
23	125,133,904	rs5933374	1.6E-05	3.9E-03	1.9E-02	187,478	Pass	WDR40C			WDR40C
23	125,321,382	rs5930842	1.1E-06	5.5E-06	4.7E-05	395,336	Pass	Olfactory Receptor			Olfactory Receptor
23	125,716,718	rs7879025	1.6E-05	7.4E-06	1.1E-03	1,000,041	Pass	Intergenic			Intergenic
23	126,716,759	rs12394158	6.9E-04	2.3E-03	4.1E-03	1,356,904					
23	128,073,663	rs7882525	5.1E-03	2.9E-02	4.6E-04	3,427			3,427		LOC392539
23	128,077,090	rs763301	5.1E-05	3.3E-03	1.7E-05	115,636	Pass	LOC392539			LOC392539
23	128,192,726	rs2103520	1.7E-05	2.3E-04	5.3E-04	418,522	Pass	Intergenic			Intergenic
23	128,611,248	rs909657	3.2E-05	1.7E-03	2.1E-04	1,647,752	Pass	APLN			APLN
23	130,259,000	rs6529475	1.2E-04	2.1E-04	9.2E-04	499,424					
23	130,758,424	rs5977554	4.7E-04	4.1E-03	9.9E-04	3,538,940					
23	134,297,364	rs12689508	4.2E-03	9.2E-04	9.1E-03	312,668					
23	134,610,032	rs5975558	1.5E-04	7.8E-03	2.6E-02	768,678					
23	135,378,710	rs1203592	1.5E-05	8.9E-04	5.8E-03	1,234,749	Pass	BRS3			BRS3
23	136,613,459	rs2366299	1.7E-06	6.4E-04	1.2E-03	509,201	Pass	LOC389892			LOC389892
23	137,122,660	rs677119	4.8E-05	5.1E-03	2.6E-04	2,747,331	Pass	Intergenic			Intergenic
23	139,869,991	rs11095848	4.2E-06	1.9E-04	1.1E-03	233,160	Pass	SPANXB1			SPANXB1
23	140,103,151	rs5907830	9.9E-04	1.5E-02	5.4E-02	594,459					
23	140,697,610	rs5908046	7.5E-05	1.4E-02	4.4E-02	291,098	Pass	Intergenic			Intergenic
23	140,988,708	rs5907221	3.4E-06	9.1E-05	1.6E-04	190,936	Pass	Intergenic			Intergenic
23	141,179,644	rs4824895	4.6E-04	1.8E-02	3.2E-03	144,480					
23	141,324,124	rs766118	6.0E-06	1.9E-01	5.7E-02	1,086,953	Pass	Intergenic			Intergenic
23	142,411,077	rs16979343	3.2E-04	1.6E-03	3.8E-04	164,918					
23	142,575,995	rs12557229	1.3E-04	2.8E-02	7.6E-01	7,651					
23	142,583,646	rs17259181	1.4E-04	3.7E-02	6.9E-01	1,010,504					
23	143,594,150	rs11094321	6.1E-06	3.3E-04	2.1E-02	1,437,140	Pass	Intergenic			Intergenic
23	145,031,290	rs5920016	3.0E-04	2.1E-02	7.1E-02	162,626					
23	145,193,916	rs1339482	4.7E-04	4.9E-02	2.8E-02	412,752					
23	145,606,668	rs409263	4.8E-04	1.0E-03	4.1E-03	555,375					
23	146,162,043	rs5905031	8.1E-05	1.2E-02	4.9E-03	234,297	Pass	MIRN510			MIRN510
23	146,396,340	rs524558	6.3E-04	1.7E-03	3.1E-02	528,831					
23	146,925,171	rs5904849	7.6E-04	7.7E-04	2.9E-02	209,797					
23	147,134,968	rs5936450	5.7E-05	5.0E-04	2.3E-04	158,645	Pass	Intergenic			Intergenic
23	147,293,613	rs5936352	5.2E-05	8.4E-04	9.0E-03	194,407	Pass	LOC347509			LOC347509
23	147,488,020	rs5980375	3.1E-08	6.9E-05	4.7E-04	10,889	Pass	AFF2	10,889		AFF2
23	147,498,909	rs7881705	3.0E-09	1.7E-05	4.6E-05	388,892	Pass	AFF2			AFF2
23	147,887,801	rs6641482	4.0E-04	3.4E-03	8.7E-03	224,162					
23	148,111,963	rs5980419	1.4E-05	5.4E-03	5.6E-03	208,770	Pass	Intergenic			Intergenic
23	148,320,733	rs17371	8.1E-05	2.2E-03	4.5E-04	1,266,787	Pass	IDS			IDS
23	149,587,520	rs10521889	1.3E-05	2.9E-05	7.2E-05	1,068,950	Pass	MTM1			MTM1
23	150,656,470	rs5970100	1.8E-04	1.2E-02	1.0E-01	1,082,139					
23	151,738,609	rs6627265	4.8E-04	4.3E-05	1.1E-03	418,541	Pass	CETN2			CETN2
23	152,157,150	rs1894344	7.6E-08	4.1E-05	4.4E-05	228,318	Pass	MAGEA1			MAGEA1
23	152,385,468	rs5945399	2.5E-05	3.7E-04	1.9E-03	92,518	Pass	UCHL5IP			UCHL5IP
23	152,477,986	rs2980017	9.7E-05	4.7E-03	8.0E-02	92,824	Poor	ATP2B3			ATP2B3
23	152,570,810	rs6571284	6.8E-06	8.0E-04	2.3E-02	633,767	Pass	DUSP9			DUSP9
23	153,204,577	rs5987245	7.7E-05	1.2E-03	1.5E-02		Pass	TKTL1			TKTL1
24	735,289	rs6579619	5.7E-03	2.1E-04	1.2E-02	1,745,832					
24	2,481,121	rs5982552	3.3E-04	8.4E-03	5.0E-03						

wellcome trust

Case-Control Consortium

Application for Access to Genotype Data

Name of applicant and co-applicant(s), including affiliations and contact details.
Please ensure that both telephone numbers and email addresses are included.

Massimo Trucco, MD (Applicant)
Children's Hospital of Pittsburgh
Pittsburgh, PA 15213 USA
TEL: (412) 692-6570
FAX: (412) 692-5809
E-mail: mnt@pitt.edu

Bernie Devlin, Ph.D. (Co-Applicant)
Department of Psychiatry
University of Pittsburgh School of Medicine
Pittsburgh, PA 15213 USA
TEL: (412) 246-6642
FAX: (412) 624-8181
E-mail: devlinbj@upmc.edu

Kathryn Roeder, Ph.D. (Co-Applicant)
Department of Statistics
Carnegie Mellon University
Pittsburgh, PA 15213
TEL: (412) 268-2513
FAX: (412) 268-7828
E-mail: roeder@stat.cmu.edu

Title of Project
In less than 30 words.

DETECTION OF SUSCEPTIBILITY LOCI IN MULTIFACTORIAL DISEASE

Genotype Data Requested
Please indicate which disease and/or control genotypes you require.

Type 1 Diabetes
Type 2 Diabetes
Bipolar Disorder
Rheumatoid Arthritis
Coronary Heart Disease
Hypertension
Inflammatory Bowel Disease

1958 British Birth Cohort controls
UK Blood Service controls
Quantile Normalised Data for Control and Disease Samples

Research Question

Please give a brief description of the project and its specific aims in no more than 500 words. This should include specific details of what you plan to do with the data and include key references.

As platforms for genome-wide association studies (GWAS) become standardized, numerous resources will become available. Analyses of GWAS are enhanced by the availability of large, well characterized, populations such as the WTCCC cohort. How to exploit these datasets effectively is an open question. As part of our work into the genetics of Type 1 Diabetes (T1D) we have developed a method to match, by genetic ancestry, controls to cases (1). Applying the method to the GWAS obtained from our cohort of T1D (N=394) and non-T1D (N=2,144) participants we have shown that GEnetic Matching (GEM) resulted in a reduced rate of false positives as well as highlighting of true positive signals for association.

The goal is to exploit our genetic data by combining it with the WTCCC resource while developing approaches that improve power of association studies to identify true signals. We are proceeding with a staged analysis for T1D using an independent cohort of cases (N=1,300) and controls (N=2,500) recruited among European-Americans and residents from Germany. Development of GEM as a method for elucidating association between SNP and T1D has led us to anticipate that combining our cohort with the WTCCC dataset will result in improved power to detect genes influencing disease.

Substantial progress has been made toward the completion of our goal. In our recent publication (1) we developed GEM as a means to control for the impact of ancestry on association testing, and showed how using GEM allows multiple datasets to be combined for discovery of genetic association. We contrasted this approach with other popular methods for controlling the spurious effects of ancestral heterogeneity, namely Eigenstrat (2). Compared with Eigenstrat, the GEM algorithm improved data analysis, highlighting true associations by reducing the number of confounding signals linked to false positives. In fact, genetic matching identified confirmed T1D loci *CTLA4*, *IL2RA*, and *PTPN22*.

Our experimental plan is to use the combined results from GWAS performed in Pittsburgh and by WTCCC to optimize GEM methodology, and apply the GEM algorithm during analysis of the complete WTCCC dataset. We anticipate that, when the Pittsburgh 'common' controls are combined with the WTCCC dataset, the method will increase power to detect true association. In order to accomplish these goals we request access to the complete WTCCC dataset of common diseases and controls. Our initial focus will be to enhance the power of the GWAS performed in Pittsburgh and by WTCCC. We will then direct our efforts towards exploiting the complete WTCCC dataset delivering an analysis with increased power to detect true signals for association with T1D. This will be our primary analysis. If our results are promising, secondary analyses will target each of the disease sets in turn, using the other, appropriate data sets as controls. Finally it is possible that some of the WTCCC common diseases have etiological factors in common. For example it is plausible that T1D, rheumatoid arthritis, and inflammatory bowel disease inherit risk from a common locus. Joint analyses using GEM methods will be useful to evaluate this possibility.

1. Luca et al. On the use of general control samples for genome-wide association studies: genetic matching highlights causal variants. Am J Hum Genet (*in press*).

2. Price et al. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006 38:904-909.

Feasibilty

Please describe your experience and expertise, and that of your collaborators, and how this will be applied to the proposed study. Relevant publications should be cited.

Massimo Trucco, MD (Applicant) is Chief of the Division of Immunogenetics at Children's Hospital of Pittsburgh, Hillman Professor of Pediatric Immunology and professor of Pediatrics at the University of Pittsburgh School of Medicine. Dr. Trucco's research is focused on the genetic and immunologic mechanisms underlying susceptibility to Type 1 Diabetes.

Bernie Devlin, Ph.D. (Co-Applicant) is Associate Professor of Psychiatry and Human Genetics and Director of the Western Psychiatric Institute and Clinic Computational Genetics Program. Dr. Devlin's research is focused on developing statistical approaches for examining genetic polymorphisms associated with human disease. He has pioneered methods for developing predictive models of linkage disequilibrium and genomic structure as well as novel analytical solutions for identifying positive genetic signatures from genome-wide analyses.

Kathryn Roeder, Ph.D. (Co-Applicant) is Professor of Statistics at Carnegie Mellon University. Dr. Roeder's research is focused on population genetic and statistical issues surrounding how best to identify disease associated genomic variants. Dr. Roeder has been awarded a National Young Investigator Prize from the National Science Foundation and has been the recipient of the Myrto Lefkopoulou Distinguished Lectureship at the Harvard School of Public Health. She is currently an associated editor for the Journal of the American Statistical Association and the American Journal of Human Genetics.

Relevant Publications (Applicant and Co-Applicant are indicated in bold font):

Devlin B, **Roeder K**, Wasserman L. (2001) Genomic control, a new approach to genetic-based association studies. Theor Popul Biol 60:155-166.

Devlin B, **Roeder K**, Bacanu SA. (2001) Unbiased methods for population-based association studies. Genet Epidemiol 21:273-284.

Luca D, Ringquist S, Klei L, Lee AB, Gieger C, Wichmann H-E, Schreiber S, Krawczak M, Lu Y, Styche A, **Devlin B**, **Roeder K**, **Trucco M**. (in press) On the use of general control samples for genome-wide association studies: genetic matching highlights causal variants. Am J Hum Genet.

Morel PA, Dorman JS, Todd JA, McDevitt HO, **Trucco M**. (1988) Aspartic acid at position 57 of the HLA-DQ beta chain protects against type I diabetes: a family study. Proc Natl Acad Sci USA 85:8111-8115.

I have read and agree to abide by the terms and conditions outlined in the Data Access Agreement and the Publication Policy?

Yes	☒
No	☐

Table of Contents

American Society for Histocompatibility and Immunogenetics

Features

Highlights

Departments

Editor's Note

Individuals interested in submitting articles for the *ASHI Quarterly* should observe the following requirements (see p.38 for additional details):

- All articles must be submitted via e-mail in Microsoft Word format
- All articles must be double-spaced

Article submissions should be forwarded to:

Editor-in-Chief
Francesco Marincola
E-mail: fmarincola@cc.nih.gov

Copy deadlines for future issues:

Fourth Quarter 2007: November 5, 2007

ASHI Quarterly

Volume 31, Number 3

ASHI Quarterly is published quarterly by the American Society of Histocompatibility and Immunogenetics, 15000 Commerce Parkway, Mount Laurel, New Jersey 08054. Subscription rates are $35 within the United States and Canada, $40 outside the United States and Canada. Subscriptions are accepted on a calendar year basis. Remittance should be made by check, draft or money order in U.S. currency, payable to the American Society for Histocompatibility and Immunogenetics. Notify ASHI of change of address at least 30 days before date of issue by sending both the old and new address. Multiple reproductions of any material in this publication is permitted only under license from the American Society of Histocompatibility and Immunogenetics with the written permission of the authors. Fair use of such material by individual readers and non-profit libraries, such as making single copies of an article for use in teaching or research, is allowed. Consent of the authors and ASHI must be obtained to reprint figures, tables or excerpts from the printed matter. The statements and opinions are not necessarily those of ASHI. ASHI does not endorse any products.

www.ashi-hla.org

A Web-Based Program for Pyrosequencing Primer Design

Steven Ringquist, Christopher Pecoraro, & Massimo Trucco
Division of Immunogenetics
Department of Pediatrics
Rangos Research Center
Children's Hospital of Pittsburgh
University of Pittsburgh School of Medicine

Address correspondence to: Massimo Trucco, MD, Division of Immunogenetics, Rangos Research Center, Children's Hospital of Pittsburgh, 3460 Fifth Avenue, Pittsburgh, PA,15213. E-mail: mnt@pitt.edu; Tel: (412) 692-6570; Fax: (412) 692-5809.

Introduction

This review presents a brief description of the Selection of Oligonucleotide Primer for PCR and Pyrosequencing (SOP3) software. The web-based application enables efficient design of polymerase chain reaction (PCR) and pyrosequencing primers that can be used for analysis of genetic polymorphisms, including single nucleotide polymorphisms (SNPs). The software automates the process of retrieving genomic DNA sequences and locations of SNPs from the human genome database, and takes into consideration specific restrictions necessary during the design of pyrosequencing assays.[1,2] The application aids in the sorting of SNPs for biologically interesting properties (e.g., the association of a SNP with a particular genetic structural element and whether it codes for an amino acid change in the resulting protein, as well as its heterozygosity within a population) that, in turn, adds to the usefulness of particular primer sets during physical mapping studies.

Pyrosequencing was developed for the analysis of expressed sequence tags (EST), enabling accurate sequencing of short stretches of DNA. This methodology has found increasing acceptance as a method for analysis of SNPs, insertion/deletion polymorphisms, and for investigating sites of DNA methylation. Sequencing of a maximun length of 40 nucleotides is commonly performed while a maximum of 150 nucleotides sequenced has been reported.[3,4] When applied to genotyping human leukocyte antigen (HLA), the method is capable of distinguishing between allelic pairs commonly found to be ambiguous using other methods [e.g., sequence specific oligonucleotide probe (SSOP) and sequence base typing (SBT) genotyping methods.][4] Pyrosequencing allows increased resolution of allelic pairs due to its ability to accurately resolve heterozygous nucleotides by enabling out-of-phase sequencing.[5] Performed by stepwise addition of dNTPs, pyrosequencing enables synthesis of the nascent nucleotide chain that is extended one nucleotide residue per dispensation step. Detection of the nucleotide sequence is performed by way of a chain of enzymatic reactions involving DNA polymerase, apyrase, ATP sulfurylase, as well as luciferase.[3,6] Each nucleotide is dispensed independently into the reaction mix. The graphic display, depicting the incorporation of a particular nucleotide, is in the form of a nucleotide dispensation event (x-axis) versus the intensity of emitted light (y-axis) and is referred to as a pyrogram (Figure 1). The pyrosequencing reaction is quantitative, in that increased light intensity is produced upon incorporation of multiple nucleotides. Thus, DNA sequence is determined by examination of light intensity emitted immediately after nucleotide dispensation and is known as sequencing by synthesis. Positions of polymorphic sequence are identified due to their exhibiting variable incorporation patterns as determined by the genotype.

Software for Pyrosequence Primer Design

SOP3 is a web-based software application (http://imgen.ccbb.pitt.edu/sop3) that has been specifically developed for use in pyrosequence assay design.[1,2] The webserver hosts an integrated biological database consisting of genetic variants as well as the DNA sequence of the human and mouse genomes, including validated human SNPs contained within the HapMap comparative genome project. The software provides a batch analysis tool that accepts the names of as many as 100 genetic loci, 100 SNP identifiers, or a 100 kb section of genomic DNA under investigation. The application returns a set of primer trios for PCR (2 primers) and pyrosequencing (1 primer) for the corresponding genetic variants listed in dbSNP. Genomic DNA sequences are stored on the web server making them readily accessible to the application. Detailed instructions for using SOP3 are available online using the Assay Design User Guide (http://imgen.ccbb.pitt.edu/sop3/entry_box.html).

The graphical interface for the SOP3 application provides an array of user-defined settings to allow customization of primer trio design. Genomic sequences from queried polymorphisms and their flanking sequences are returned to the SOP3 application's core algorithm and enable the selection of suitable primer trios. Choice of PCR primers is based on multiple criteria including the distance between PCR primer annealing sites and the SNP, as well as allowing for the selection of optimal genomic flanking sequence for annealing of sequencing primer. Additional criteria for maximizing PCR amplification yield are included. For example, the search identifies suitable primers at multiple annealing sequences on either side of the SNP. The list of candidate PCR primers is evaluated for their ability to form a competing secondary structure that may interfere with primer annealing; the

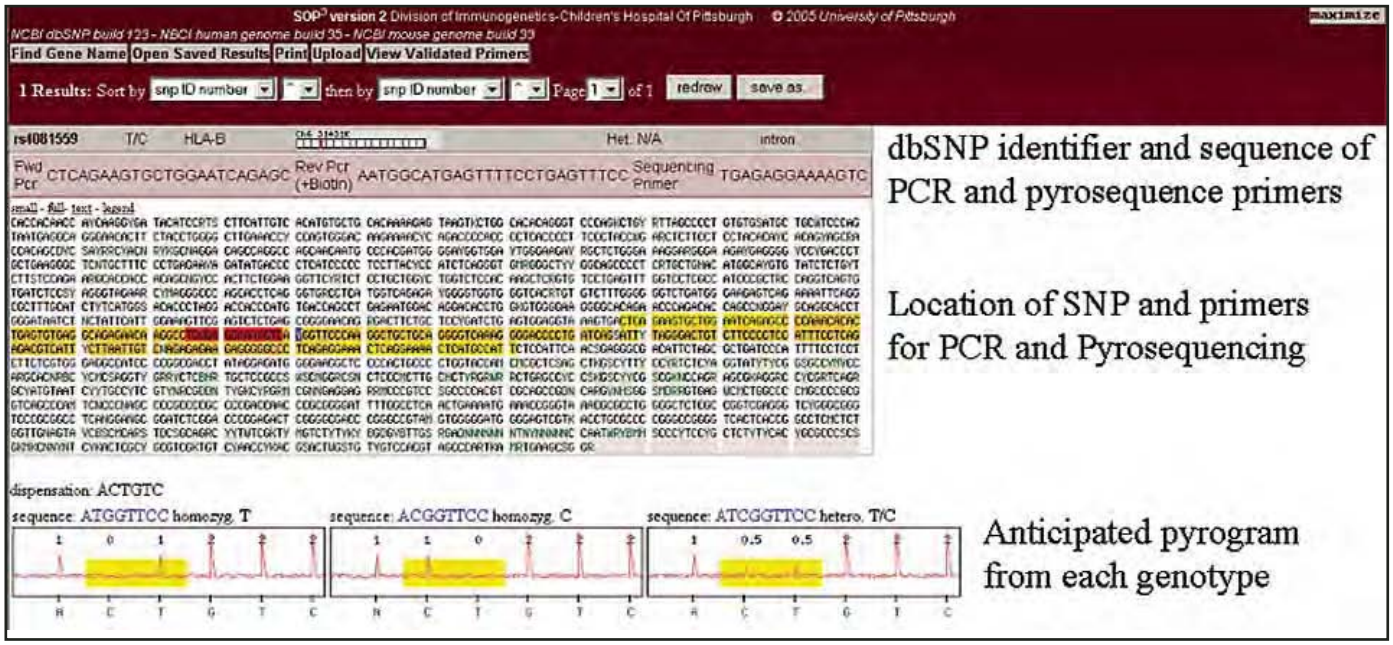

Figure 1

Results of the query for SNP rs4081559 contained within intron 2 of HLA-B as obtained from the SOP3 application. Upper Panel. The output summarizes the SNP under consideration by referencing the dbSNP identifier (i.e., rs4081559) and provides the recommended sequence of primers for PCR and pyrosequencing. Middle Panel. The location of SNP and primers are indicated in blue (query SNP), red (pyrosequencing primer), and yellow (PCR primers). The region amplified during PCR is indicated in gold. The locations of nearby SNPs are indicated (pink). Bottom Panel. The anticipated pyrograms from each genotype are shown. Pyrograms plot the nucleotide dispensation order (x-axis) versus the light emitted (y-axis) during sequencing by synthesis steps of pyrosequencing.

predicted stability of primer annealing is monitored by calculating the melting temperature (Tm); sequence complexity is measured by maintaining representative frequencies of each nucleotide residue in the final primer sequence. Identification of nearby polymorphic sites is performed in order to improve the design of primer trios for sequencing as well as to enable possible multiplexing of assays for genotyping these residues.

The software generates lists of candidate forward and reverse PCR primers that pass the individual tests. These oligonucleotide sequences are then compared in order to obtain compatible primer pairs for use during PCR amplification. This is accomplished by analyzing primers for regions of reverse complementary sequence that may initiate primer dimer amplification. Final optimization of PCR primers is accomplished by modifying the PCR forward primer by addition of an extra-genomic nucleotide motif to the 5'-end. This results in suppression of competing sequencing reactions during genotyping analysis that may occur whenever the template strand 3'-end is capable of annealing to an internal sequence motif, thus repressing initiation of intra-molecular sequencing.[7] The resulting PCR primers are returned to the user as oligonucleotide sequences recommended for laboratory testing which are either presented graphically or in a tab-delimited format (Figure 1, top panel).

Computer software for designing primers for use during polymerase chain reaction (PCR) is available from a number of sources (Table 1). A principle use of these applications is to aid assay design when evaluating the association between phenotypic changes and the inheritance of genetic markers, such as single nucleotide polymorphisms (SNPs) and nucleotide insertions and deletions. Commercial software is also available for generating PCR and sequencing primers (Table 1). For example, the manufacturer of the pyrosequence instrument Biotage LLC provides assay design software aimed at developing primer sets suitable for use during pyrosequencing. These applications provide access to PCR and sequencing primer design algorithms but require the manual input of sequence along with identification of the polymorphic residues to be evaluated. For genetic studies where many polymorphisms are being investigated, such as in the investigation of a set of candidate markers for association with a disease, this requirement creates a bottleneck during assay development. Thus, improved approaches to assay design that provide integrated genomic sequence, genetic variant, and functional information increase the efficiency by which genotyping assays can be developed and validated by laboratory testing.[8]

The objective of the SOP3 application has been to provide a method to effectively search genomic databases for SNPs and insertion/deletion polymorphisms in order to design genotyping assays. The application has been used to design primers that have been validated for a variety of SNPs within loci suspected of correlating with risk towards developing disease as well as HLA genotyping.[9-12] Localizing the NCBI dbSNP database and the data from University of California Santa Cruz (UCSC) GoldenPath for the human genome allowed the creation of structured query language (SQL) to present genomic variation data as well as annotation of functional genomic elements in a customized format. Using the SOP3 application for a multi-loci genetic marker association study is aided by the use of the integrated Entrez Gene annotation, which identifies the start and end markers of a gene, as well as the start and end markers of exons, introns, and the location of polymorphisms associated with splice site regions. SOP3 designed primer trios continue to be evaluated

Website	Web Address	Description
Primer Design Software:		
Integrated DNA Technologies	http://www.idtdna.com	Supplier of oligonucleotides. The website also contains an extensive selection of scientific tools for PCR primer design.
Primer3	http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi	Primer3 is a widely used program for designing primers for PCR.
SOP3	http://imgen.ccbb.pitt.edu/sop3	Web-based software specifically tailored for designing primers for pyrosequencing.
Human Genome Resources:		
NCBI	http://www.ncbi.nlm.nih.gov/	The National Center for Biotechnology Information (NCBI) is a national resource for molecular biology information containing public databases and software tools for analyzing genome data.
dbSNP	http://www.ncbi.nlm.nih.gov/SNP/	The dbSNP database serves as a central repository for both single base nucleotide subsitutions and short deletion and insertion polymorphisms.
Entrez Gene	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene	Entrez Gene is a searchable database of genes maintained by the National Center for Biotechnology Information.
HapMap	http://www.hapmap.org/	The HapMap Project is an effort to identify and catalog genetic similarities and differences in human beings.
Pyrosequencing Resources:		
Biotage, AB	http://www.biotage.com	Leading supplier of pyrosequencing reagents & instruments.

Table 1. Websites and Internet Resources Related to Pyrosequencing.

in the laboratory as part of ongoing research into the genetics of complex diseases using human cohorts as well as mouse model systems.[4,13]

Summary

The ability to generate high resolution sequence data and throughput of hundreds of sequencing reactions per day make pyrosequencing an important method for identification of known genetic markers. Based on finding a short sequence in each allele of interest that occurs only once within the PCR generated fragment, the pyrosequencing protocol allowed rapid identification of alleles for analysis of genetic predisposition to autoimmune disease as well as for histocompatibility typing. Pyrosequencing is an ideal tool for rapid and accurate identification of genotypes in studies of the genetic association of certain HLA alleles with the likelihood of developing autoimmune disease, as well as during clinical genotyping of patients and donors for histocompatibility matching during transplantation.

Acknowledgements

The authors thank Alexis Styche for a critical reading of the manuscript. This work was supported by grants from the NIH-NLM, U19-AI056374-01 Autoimmunity Centers of Excellence (SR and MT), RO1DK24021(MT) from the National Institutes of Health, and ERHS #00021010 (SR and MT) from the Department of Defense.

References

1. Ringquist S, Pecoraro C, Gilchrist MS, Styche A, Rudert WA, Benos PV, Trucco M. SOP3v2: web-based selection of oligonucleotide primer trios for genotyping of human and mouse polymorphisms. Nucl Acids Res 2005;33:W548-W552.

2. Ringquist S, Pecoraro C, Lu Y, Styche A, Rudert WA, Benos PV, Trucco M. Web-based primer design software for genome scale SNP mapping by pyrosequencing. In: Methods in Molecular Biology, vol 373: Pyrosequencing Protocols. ed. S. March. Humana Press Inc., Totowa, New Jersey (2006).

3. Gharizadeh B, Nordstrom T, Ahmadian A, Ronaghi M, Nyren P. Long-read pyrosequencing using pure 2'-deoxyadenosine-5'-O'-(1-thiotriphosphate) sp-isomer. Anal Biochem 2002;301:82-90.

4. Ringquist S, Alexander AM, Styche A, Pecoraro C, Rudert WA, Trucco M. HLA class II DRB high resolution genotyping by pyrosequencing: comparison of group specific PCR and pyrosequencing primers. Hum Immunol 2004;65:163-174.

5. Garcia CA, Ahmadian A, Gharizadeh B, Lundeberg J, Ronaghi M, Nyren P. Mutation detection by pyrosequencing: sequencing of exons 5-8 of the p53 tumor suppressor gene. Gene 2000;253:249-257.

6. Ronaghi M, Karamohamed S, Pettersson B, Uhlen M, Nyren P. Real-time DNA sequencing using detection of pyrophosphate release. Anal Biochem 1996;242:84-89.

7. Ronaghi M, Pettersson B, Uhlen M, Nyren P. PCR-introduced loop structure as primer in DNA sequencing. Biotechniques 1998;25:876-878.

8. Stein LD. Integrating biological databases. Nature Rev Genet 2003;4:337-345.

9. Todd JA, Bell JI, McDevitt HO. HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 1987;329:599-604.

10. Morel PA, Dorman JS, Todd JA, McDevitt HO, Trucco M. Aspartic acid at position 57 of the HLA-DQ beta chain protects against type I diabetes: a family study. Proc Natl Acad Sci USA 1988;85:8111-8115.

11. Sollid LM. Celiac disease: dissecting a complex inflammatory disorder. Nat Rev Immunol 2002;2:647-655.

12. Zanelli E, Breedveld FC, de Vries RR. HLA class II association with rheumatoid arthritis: facts and interpretations. Hum Immunol 2000;61:1254-1261.

13. Mathews CE, Leiter EH, Spirina O, Bykhovskaya Y, Gusdon AM, Ringquist S, Fischel-Ghodsian N. mt-Nd2 Allele of the ALR/Lt mouse confers resistance against both chemically induced and autoimmune diabetes. Diabetologia 2005;48:261-267.

3

Web-Based Primer Design Software for Genome-Scale Genotyping by Pyrosequencing®

Steven Ringquist, Christopher Pecoraro, Ying Lu, Alexis Styche, William A. Rudert, Panayiotis V. Benos, and Massimo Trucco

Summary

Design of locus-specific primers for use during genetic analysis requires combining information from multiple sources and can be a time-consuming process when validating large numbers of assays. Data warehousing of genomic DNA sequences and genetic variations when coupled with software applications for optimizing the generation of locus-specific primers can increase the efficiency of assay development. Selection of oligonucleotide primers for PCR and Pyrosequencing® (SOP^3) software allows user-directed queries of warehoused data collected from the human and mouse genome sequencing projects. The software automates collection of DNA sequence flanking single-nucleotide polymorphisms (SNPs) as well as the incorporation of locus-associated functional information, such as whether the SNP occurs in an exon, intron, or untranslated region. SOP^3 software accepts three types of user-directed input consisting of gene locus symbols, SNP reference sequence numbers, or chromosomal physical location. For human polymorphisms, SOP^3 incorporates haplotype, ethnicity, and SNP validation attributes. The output is a list of oligonucleotide primers recommended for Pyrosequencing-based typing of genetic variations. SOP^3 is available at the Division of Immunogenetics computational server found at http://imgen.ccbb.pitt.edu.

Key Words: Bioinformatics; genetics; genomics; genotyping; sequencing.

1. Introduction

Computer software for designing primers for use during PCR are available from a number of sources ***(1–3)***. A principle use of these applications is to aid assay design when evaluating the association between phenotypical changes and the inheritance of genetic markers, such as single-nucleotide polymorphisms (SNPs) including nucleotide insertions and deletions. Commercial software is also available for generating PCR and sequencing primers and for use

From: *Methods in Molecular Biology, vol. 373: Pyrosequencing® Protocols*
Edited by: S. Marsh © Humana Press Inc., Totowa, NJ

during DNA sequencing. For example, the manufacturer of the Pyrosequencing® instrument Biotage, LLC provides assay design software aimed at developing primer sets suitable for use during Pyrosequencing. These applications provide access to PCR and sequencing primer design algorithms, but require user input of FASTA-formatted sequences along with identification of polymorphic residues that are to be evaluated. For genetic studies where many polymorphisms are being investigated, such as is necessary when investigating a set of candidate markers for association with a disease phenotype, this requirement creates a bottleneck during assay development. Improved approaches to assay design that provide integrated genomic sequence, genetic variant, and functional information increase the efficiency by which genotyping assays can be developed and validated by laboratory testing ***(4)***.

SOP^3 is a web-based software application that has been developed for use in Pyrosequencing assay design and is available at http://imgen.ccbb.pitt.edu ***(5)***. The computational webserver warehouses an integrated biological database comprised of genetic variants, as well as DNA sequence of the human and mouse genomes, including validation of human SNPs within the HapMap comparative genome project. The software provides a batch analysis tool that accepts the names of as many as 100 genetic loci, 100 SNP identifiers, or a 100-kb section of genomic DNA under investigation. The application returns a set of primers for PCR and Pyrosequencing for all corresponding genetic variants listed in dbSNP along with flanking genomic DNA sequences contained within the human and mouse genome resources. Local warehousing of genomic sequences, making them accessible to the application, enables the software to provide the choice of locus-specific primer trios for PCR and Pyrosequencing for use in rapid genotype analysis.

2. Materials

2.1. System Configuration

1. The SOP^3 application, genomic sequence, and genetic variation data from the human and mouse genome projects are warehoused on a computational webserver purchased from @Xi Computer (San Clemente, CA). The software application is written in preprocessor hypertext protocol v5.0.3 and is associated with a MySQL 4.1 database developed on a Linux SuSE Enterprise Server 8 for the AMD64 operating system with Apache web server v2.0.48. The computational server consists of dual AMD Opteron 246 64-bit processors with 1024 KB Cache, 8192 MB random access memory, and four 250 GB hard drives.

2.2. Connecting to the Webserver

1. The SOP^3 website can be accessed over the internet at http://imgen.ccbb.pitt.edu. The application works best when viewed in Internet Explorer v6 or higher but can

also be viewed in Firefox as well as Netscape. Internet browsers are available for download from http://www.microsoft.com/windows/ie, http://browser.netscape.com, and http://www.mozilla.org, respectively.

3. Methods

The SOP[3] application has been designed for use with genotyping protocols and enables locus-specific PCR amplification of candidate SNPs for sequence analysis. Filtering of genetic variations for validated and HapMap-associated SNPs improves the efficiency of assay development for studying the role of DNA polymorphisms in determining the phenotype in models of genetic inheritance **(Table 1)**. An advantage of the SOP[3] application is that design of genotyping studies can be developed using the assay design algorithm by filtering all available genetic variants for those specified by user-directed query, such as specifying human or mouse and using only validated nucleotide polymorphisms (*see* **Note 1**). The application accepts multiple queries of gene loci, SNPs, as well as chromosomal regions for analysis. SOP[3] provides multiple choices for selecting the most suitable genetic markers for evaluation during genotyping. The User Guide for operation of SOP[3] software is available online at http://imgen.ccbb.pitt.edu/sop3/userguide.

TABLE 1

3.1. Basic Functions

1. Searching human or mouse genetic variations. The web interface to the SOP[3] application allows the user to specify whether to search for SNPs within the human or mouse genome. The default setting is for human genetic variants and is indicated on the website by the filled in button adjacent to the label for "Human." To specify searching of the mouse database the user should click the button next to "Mouse" provided on the website (**Fig. 1**).
2. Searching by locus name. To query the SOP[3] database for SNPs associated with up to 100 locus symbols the user can enter the symbol or a list of locus symbols into the textbox. Locus symbols can be typed directly into the text box, entered using the copy and paste function on the user's computer, or by uploading from a text file. Queries are restricted to no more than 100 entries per batch. Use the check boxes provided to specify whether the query is for human or mouse genomes. Click the "Search" button to activate the application. Results will appear at the bottom of the page. Selecting the "Clear" button will refresh the application to the default settings.
3. Searching by SNP identifier. To query the SOP[3] database by SNP identifier the user can enter either a single reference sequence number or a list of polymorphism reference sequence numbers. Queries are restricted to no more than 100 entries per batch. Use the check boxes provided to specify whether the query is for human or mouse genomes. Click the "Search" button to activate the application. Results will appear at the bottom of the page. Selecting the "Clear" button will refresh the application to the default settings.

FIG 1

Table 1
Classification of Human and Mouse Genetic Variants

Classification	Number of human SNPs	Number of mouse SNPs	Description
Validation status			
HapMap	35,126	NA	The genetic variant has been genotyped by the International HapMap Consortium and incorporated into the dbSNP resource.
Validated	384,831	464,137	Multiple submissions of the genetic variant has been reported within the dbSNP resource.
Function type			
Coding: syn. unknown	34	9	The genetic variant is within the coding region of a gene. The location of the polymorphism cannot be resolved because of an error in the alignment of the exon.
Intron	3,704,388	213,968	The variant is in the intron of a gene but not within two residues from the intron/exon boundry.
Nonsynonymous change	61,963	5478	The genetic variant is nonsynonymous for the codon within the gene. This class of allele is defined by a substitution and transla tion of the allele into the codon that results in a change to the amino acid specified by the exon reading frame.
No RNA_acc / Protein_acc	357,606	30,373	The polymorphism is within 2000 nucleotides 5' or 500 nucleotides 3' of a gene feature. The variant is not in the transcript for the gene.
Splice site	971	23	The variant is in the first two or last two nucleotides on the intron of a gene.
Synonymous change	48,702	9548	The genetic variant is synonymous for the codon within the gene. This class of allele is defined by a substitution and translation of the allele into the codon that results in no change to the amino acid specified by the exon reading frame.
Untranslated region	668,282	43,414	The nucleotide polymorphism is in the transcript of a gene but not in the coding region of the mRNA.

Geographic population

Cent. Asia	46	NA	Samples from Russia and its satellite republics and from nations bordering the Indian Ocean between East Asia and the Persian Gulf regions.
Cent/S. Africa	30	NA	Samples from nations south of the Equator, Madagascar, and neighboring island nations.
Cent/S. America	11	NA	Samples from Mainland Central and South America and island nations of the western Atlantic, Gulf of Mexico, and Eastern Pacific.
E. Asia	7447	NA	Samples from eastern and south eastern Mainland Asia and from Northern Pacific island nations.
Europe	535	NA	Samples from Europe north and west of Caucasus Mountains, Scandinavia, and Atlantic islands.
Multinations	13,004	NA	Samples that were designated to maximize measures of heteroge neity or sample human diversity in a global fashion.
N. America	11,438	NA	All samples north of the Tropic of Cancer, including defined samples of US caucasians, African Americans, Hispanic Americans, and the National Human Genome Research Initiative (NHGRI) poly morphism discovery resource.
N.E. Africa and Mid. East	27	NA	Samples collected from North Africa (including the Sahara desert), East Africa (south to the Equator), Levant, and the Persian Gulf.
Pacific	144	NA	Samples from Australia, New Zealand, Central and Southern Pacific Islands, and Southeast Asian peninsular/island nations.
Unknown	479	NA	Samples with unknown geographic provinces that are not global in nature.
W. Africa	174	NA	Sub-Saharan nations bordering the Atlantic north of the Congo River and central/southern Atlantic nations.

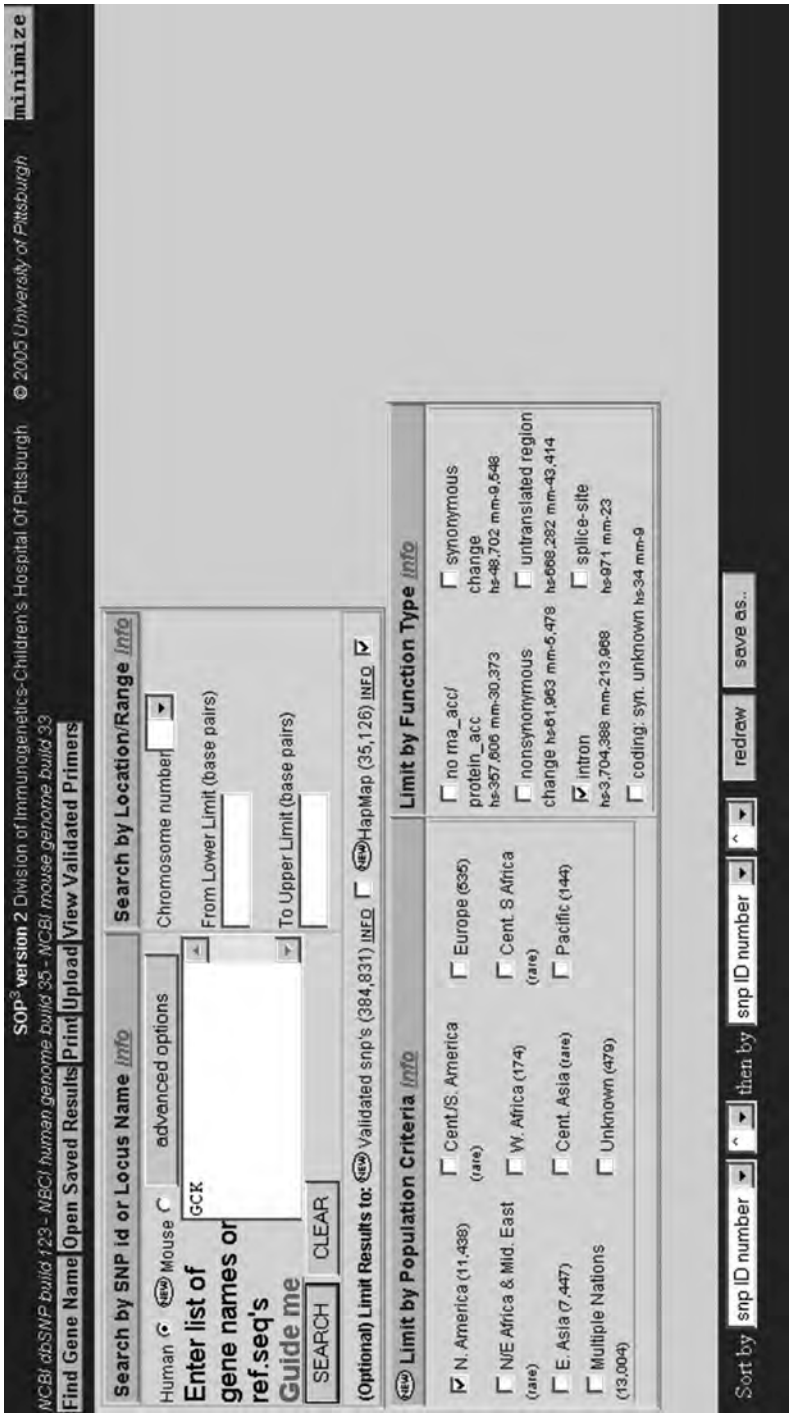

Fig. 1. User interface to the SOP[3] web-based software application. The example illustrates a search for polymorphisms associated with the human glucokinase locus, GCK. The query is limited to single-nucleotide polymorphisms linked to the HapMap project, filtered for genetic variants found in populations within North America, and for polymorphisms associated with intron elements of the RNA transcript.

AU: Is GCK referring to a gene in Figs. 1 and 2?

Table 2
Maximum Allowable Upper Limit for Screening Chromosome Range

Human chromosome	Maximum range	Mouse chromosome	Maximum range
1	245,442,500	1	195,198,653
2	242,818,021	2	181,685,801
3	199,450,386	3	160,571,871
4	191,400,945	4	154,132,574
5	180,837,593	5	149,217,787
6	170,972,421	6	149,553,910
7	206,556,958	7	133,031,105
8	146,272,185	8	128,674,990
9	138,428,984	9	124,140,960
10	135,412,916	10	130,567,357
11	134,451,003	11	121,607,694
12	132,389,146	12	114,933,529
13	114,127,336	13	116,456,691
14	106,360,250	14	117,011,917
15	100,337,960	15	104,102,711
16	88,821,548	16	98,800,952
17	80,652,345	17	93,538,550
18	76,116,152	18	90,878,013
19	63,806,020	19	60,667,351
20	62,435,629	X	160,070,598
21	46,943,948	Y	47,759,179
22	49,534,318		
X	154,823,225		
Y	57,700,652		

Values for maximum range were determined using the highest chromosomal position number associated with a SNP on the human or mouse genomes using build 35 or 33, respectively.

4. Searching by chromosomal location and range. To query the SOP[3] database by chromosomal range use the check boxes to specify whether the search will focus on human or mouse genomes. Next, use the "Chromosome Number" dropdown menu to specify which chromosome is to be used. Enter lower and upper limits for the nucleotide positions that should be screened. Queries are restricted to regions no greater than 100,000 nucleotides. The maximum allowable upper limit for screening a range of nucleotides on each chromosome corresponds to the length of the chromosome and is indicated in **Table 2**. Click the "Search" button to activate the application. Results will appear at the bottom of the page. Selecting the "Clear" button will refresh the application to the default settings.
5. Filtering the results to show only validated SNPs. Selection of the "Validated SNPs" search filter directs the application to screen genetic polymorphisms for inclusion in the list of validated polymorphisms provided by dbSNP. There are

Job: Marsh 645_8
Chapter: 03_Truc
Temp: MiMB_6x9_TempWS_place
Compositor: MBL
Date: 06.06.06
Revision: First Pass

currently 384,831 human and 464,137 mouse polymorphisms associated with this attribute. Validation status is defined as a genetic variant that has been reported by multiple submissions, is linked to frequency data for the SNP in a population, or was submitter validated by an updated submission from the original report.

6. Filtering the results to show only HapMap-genotyped SNPs. Selection of the "HapMap" search filter directs the application to screen genetic polymorphisms for those validated polymorphisms that are also included in the International HapMap Consortium haplotyping mapping project incorporated into the dbSNP database. There are currently 35,126 human polymorphisms associated with this attribute.
7. Filtering the results by limiting the population criteria. Selection of the "Limit By Population Criteria" filter directs the application to screen genetic polymorphisms by geographic origin of the sample. Selection of more than one population filter will return those polymorphisms that meet either of the criteria (i.e., a Boolean OR operation). The number of polymorphisms associated with each population group is indicated on the application's homepage.
8. Limiting the search by function type. Genetic variants may be filtered by whether they are near a gene (locus region), in a UTR (untranslated region), in an intron (intron), or in a splice site (splice site). If the variation is in a coding region, then the functional class of the variation depends on how each allele may affect the translated peptide sequence, e.g., nonsynonymous change or synonymous change. Selection of more than one of the function-type filters will return those polymorphisms that meet either of the criteria (i.e., a Boolean OR operator). A definition of these filters as well as the number of polymorphisms associated with each function type, for human and mouse, is indicated in **Table 1**.
9. Initiating design of primers for PCR and Pyrosequencing. Selecting the "Search" button will activate the application. Locus-specific primers for PCR and Pyrosequencing will be returned for those SNPs that meet the filtering criteria selected in the user input section of the web interface.
10. Reporting the results. After submitting a search the application will present the results at the bottom frame of the website. Each report consists of four sections, attribute bar, candidate primer bar, DNA sequence flanking the polymorphism, and computer-simulated Pyrosequencing data for heterozygous and each homozygous genotype (**Fig. 2**).

FIG 2

11. The locus attribute bar provides information linked to the polymorphism consisting of reference sequence number, identity of the polymorphism, locus symbol, physical location on the chromosome, heterozygosity value, and function attribute. If the filters for "Validated SNPs," "HapMap," or "Limit By Population Criteria" were selected that information will also appear in the attribute bar.
12. The candidate primer bar contains the oligonucleotide sequence (written 5' to 3') for the candidate PCR forward and reverse primers as well as the Pyrosequencing primer (*see* **Note 2**). The nomenclature used for designating a forward PCR primer is that it is on the same side of the SNP as the Pyrosequencing primer. The nomenclature for designating a reverse PCR primer is that it is on the opposite side of the SNP as the Pyrosequencing primer. The reverse PCR primer should be synthesized with a 5' biotin-TEG modification during the Pyrosequencing protocol.

AU: Pls spell out TEG, if appropriate.

Job: Marsh 645_8
Chapter: 03_Truc
Temp: MiMB_6x9_TempWS_place

Compositor: MBL
Date: 06.06.06
Revision: First Pass

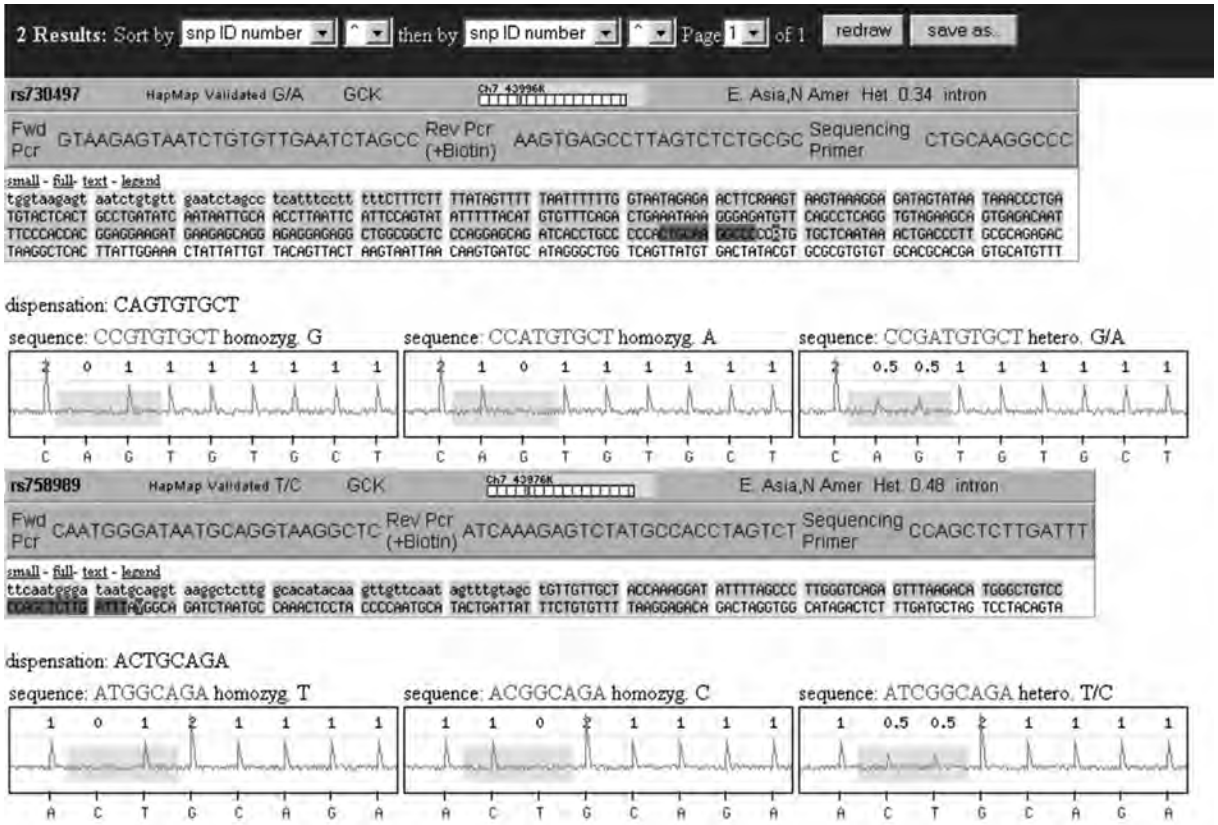

Fig. 2. Results of the query for single-nucleotide polymorphisms (SNPs) linked to the human glucokinase locus, GCK. There are two entries in SOP[3] v2 corresponding to the query. The attribute bar, candidate primer bar, DNA sequence, and simulated Pyrosequencing® data are shown. The color code in the DNA sequence box indicates the location of the PCR primers (yellow), Pyrosequencing primer (red), region of amplified DNA (orange), and location of the SNP (blue).

AU: Is GCK referring to a gene in Figs. 1 and 2?

AU: Pls modify Fig. 2 caption; there is to be no color.

13. The DNA sequence flanking the SNP is indicated in the third panel of the results frame. The user may view the "Full" sequence that was evaluated, consisting of the polymorphism and 1000 flanking residues. The DNA sequence may also be viewed in FASTA format by selecting the "Text" button. The option for viewing the flanking sequence as "Small" will display only the sequence amplified by the locus-specific PCR primers (*see* **Note 3**). The color code in the bottom frame shows PCR primers indicated in yellow, Pyrosequencing primer in red, the SNP is colored blue, the amplified region in orange, and nearby polymorphisms shown in gray. Lowercase characters indicate the location of repeat masker-identified genomic regions. A "Legend" button providing a key to the color coding is also provided.
14. Simulated Pyrosequencing data is displayed in the bottom panel of each result. A simulated Pyrogram is drawn for each genotype, i.e., heterozygous and both homozygous genotypes. A nucleotide dispensation order for evaluating these data in the Pyrosequencing software is illustrated above the Pyrogram charts. Informative dispensations for distinguishing the genotypes are indicated in yellow (*see* **Note 4**).
15. Minimization and maximization of the input/output frames can be accomplished by clicking the "Minimize/Maximize" button located in the upper right-hand corner of the web interface. A full-screen view of the results of the SOP^3 application is available, allowing easier browsing of the results frame.
16. Sorting the results by reference SNP number, heterozygosity value, physical position on the chromosome, gene name, or function can be accomplished using the "Sort By" dropdown menus. The "Sort By" function is located in the horizontal bar found along the middle of the web interface. Results may be viewed in ascending "^" or descending "v" order. Results on the web interface are presented 20 at a time. When greater than 20 results are returned they can be viewed by choosing the "Page" dropdown menu. The total number of results and pages of results is also indicated. After selection of functions for sorting and page display the results can be refreshed by clicking the "Redraw" button located on the middle toolbar.
17. Results from the SOP^3 application can be saved in an Excel formatted text file. Select this function by clicking the "Save As." This will open a text box for designating a file name and a "Save" button for uploading the results to the SOP^3 server. The saved file can be viewed by selecting the "Open Saved Results" button located at the top of the page. Saved files are listed alphabetically and each file contains a tab-delimited list of PCR and sequencing primers, as well as SNP attributes such as reference sequence number, heterozygosity value, and function attribute. Selection of the "Save All" button will allow all the results to be transferred to the web server. Unselecting of the "Save All" button will allow only the page of results presently displayed in the bottom frame to be saved.
18. Selecting the "Clear" button located in the upper frame of the web interface will refresh the application to the default settings.

3.2. Advanced Options for PCR Primer Design

1. Selection of the "Advanced Options" button opens a list of PCR and sequencing primer design settings used in the assay design algorithm. This function allows

the user to specify customized settings for primer selection that are different from the default settings. Clicking the "Advanced Options" a second time will hide these values. Selecting the "Clear" button will refresh the application to the default settings.

2. The display for PCR primer design settings allows the user to adjust the primer selection criteria.
3. "Melting Temperature (T_m) ºC" default setting is 60ºC. T_m is calculated using the relationship (T_m = 16.6*log[cation concentration] + 41*[fraction of GC] + 81.5, where the concentration of cation was estimated to be 0.1 *M*) described by Schildkraut and Lifson ***(6)***. User-directed changes to this value instruct the application to design primers to the indicated minimum T_m value.
4. The "Minimum Foldover" filter is set at a default of five. This value reflects the minimum allowable number of contiguous base-paired residues that can occur within a candidate PCR primer. The function allows user-directed input in order to minimize the formation of potential primer secondary structure that can interfere with PCR efficiency.
5. The "Unique N'mer Maximum Length" filter default setting is six and specifies the maximum number of 3'-end residues that meet the requirement of only occurring once within 1000 nucleotides flanking either side of the polymorphism.
6. The "A/T Test" filter specifies that at least one A or T residue will occur in the last three 3' nucleotides of the PCR primer.
7. The "Residue Thresholds" filter specifies the frequency range for each A, T, C, G nucleotide residue in the candidate PCR primer. The default setting for this filter is minimum 14% and maximum 40% nucleotide composition.
8. The "Flank Length (Maximum 1000)" setting default setting is 1000 and denotes the length of DNA flanking the polymorphism that is used to screen for suitable PCR primers.
9. The filter "Use Repeat Masker" indicates whether the results of RepeatMasker (repeatmasker.org) will be used when evaluating the DNA sequence flanking the SNP. When selected, residues that are identified as being included in a repetitive region will be masked off and will not be considered when making the PCR primer, thus generating primers that will not anneal to regions of repetitive DNA. Using this option will generally decrease the number of primers found, but will better ensure that primers do not anneal to multiple sites.
10. The filter for "PCR Product Size" determines the range of length of an acceptable amplified PCR product. The default setting for this filter is minimum 200 bp and maximum 500 bp.

3.3. Advanced Options for Pyrosequencing Primer Design

1. User-directed changes to the Pyrosequencing primer design settings are accomplished using the three interactive boxes available on the lower panel of the "Advanced Options" window.
2. The "Melting Temperature (T_m) ℃" default setting is 40℃. T_m is calculated using the relationship (T_m = 16.6*log[cation concentration] + 41*[fraction of GC] + 81.5, where the concentration of cation was estimated to be 0.1 *M*) described

Job: Marsh 645_8
Chapter: 03_Truc
Temp: MiMB_6x9_TempWS_place
Compositor: MBL
Date: 06.06.06
Revision: First Pass

by Schildkraut and Lifson ***(6)***. User-directed changes to this value instruct the application to design each sequencing primer to the indicated minimum T_m value.

3. The filter for "T+A Percentage" evaluates the design of the candidate sequencing primer for % AT content ensuring that this value falls within the specified range. The default range is minimum 30% and maximum 65%.
4. The interactive box for designating the "Distance From SNP" provides a tool for specifying the minimum number of bases away from the SNP to initiate the design of the Pyrosequencing primer. The default range is minimum 1 and maximum 30 nucleotides. The software algorithm is designed to choose the closest Pyrosequencing primer to the SNP but with the range of distance designated by the "Distance From SNP" function.

3.4. Exploring the Website

1. General information regarding warehoused genomic and genetic variant data is indicated in the masthead of the web interface. For example, the build of dbSNP and human and mouse DNA sequences used for v2 of SOP[3] are build 123, 35, and 33, respectively. Updates to the SOP[3] database are performed twice a year and are indicated in the masthead.
2. The "Find Gene Name" button opens a link to the search page for determining the locus symbol for either human or mouse genes as specified by the user. The "Fine Gene Name" search box allows simple Boolean searches. It allows and, or, not. It uses "*" to indicate a wildcard character and quotation marks to designate a phrase. Queries are submitted using the "Submit Query" button.
3. The "Open Saved Results" button opens a link to the list of saved files available on the SOP[3] web server. Selection of a file provides a tab-delimited document consisting of the locus symbol, reference sequence number, allele, heterzygosity value, chromosome location, primer sequences, and size of expected PCR product. The tab-delimited document can be saved to the user's computer and opened in Microsoft Excel.
4. The "Print" button provides a link to the user's print command for use in obtaining a hardcopy of the results.
5. The "Upload" button opens a "Text Box," "Browse" button, "Submit" button, and "Cancel" button for uploading to the web server a text file containing a batch of locus symbols or reference sequence numbers from a file located on the user's computer. Selection of the "Browse" button prompts the user to indicate the location of a carriage return delimited text file. Selection of the "Submit" button uploads the query batch to "Search By SNP ID Or Locus Name," i.e., the application's main entry box. The application is initiated by selecting the "Search" button. Selection of the "Cancel" button removes the upload feature from the web interface and allows the user to query genomic SNPs using the basic functions available on the website.
6. The "View Validated Primers" button opens a link to the primer trios that have been made available for general use. They are listed in a tabulated format and include reference sequence, locus, forward PCR primer, reverse PCR primer, Pyrosequencing primer, and PCR product length.

3.5. Online User Guide

1. The "Guide Me" button opens a link to the online user guide for the SOP[3] application. The user guide is updated along with changes to the application and the warehoused database (*see* **Note 5**). Explanations of the choices available when using the application are indicated in alphabetical order.
2. The "Info" buttons open information links for selected functions available on the web interface. For example, selection of the "Info" button next to the "Limit By Function Type" heading opens a new browser window to the online user guide describing the use of this filter when selecting SNPs for genetic analysis.

4. Notes

1. Software applications for primer design are available from several sources. Commonly used web-based applications for designing PCR primers are available at the Primer3 (http://www.primer3.org) and IDT Corporation (http://www.idtdna.com) websites ***(1)***. Both applications generate oligonucleotide sequences for locus-specific PCR amplification. They are limited, however, in that they do not provide a sequencing primer for Pyrosequencing. Another application, specific for Pyrosequencing, is Assay Design Software available commercially from Biotage, LLC. This application provides access to a primer design algorithm for locus-specific PCR and SNP-specific Pyrosequencing. The application, however, requires user input of each SNP and DNA flanking sequence. The Biotage Assay Design Software can be operated using a batch mode once the sequences are arranged in a text file on the user's computer. These applications and the SOP[3] software provide complimentary solutions for selecting robust primers for use during pyrosequencing-based typing of genetic variants.
2. Addition of a unique nucleotide motif to the 5'-end of the forward PCR primer occurs in order to avoid formation of secondary structure in the biotinylated template strand, which can lead to competing sequencing signal owing to self-priming during Pyrosequencing ***(7–9)***. The nucleotide motif is chosen from a list of all possible nucleotide combinations that do not occur within the DNA sequence flanking the SNP.
3. Validating PCR primers designed by the computer application should be performed by PCR amplification followed by analysis of the product with agarose gel electrophoresis. Selection of primers assumes a standard PCR condition using an annealing temperature of 60°C. It is recommended that, when possible, PCR product yield be examined at 1 and 2 m*M* $MgCl_2$ and over an annealing temperature range from 54 to 65°C using a gradient thermal cycler instrument. PCR primers are considered validated if they result in a single amplified DNA product of the expected size as visualized by agarose gel electrophoresis.
4. Validating Pyrosequencing primers is accomplished by performing the Pyrosequencing reaction using a control DNA of known quality and combined with a set of negative control Pyrosequencing reactions. The negative control reactions are designed to allow troubleshooting of the sequencing reaction by analyzing the level of background signal associated with the sequencing primer, biotinylated reverse PCR primer, and possible alternative PCR products. Recom-

mended negative control reactions are as follows: (1) PCR product without the addition of Pyrosequencing primer; (2) PCR negative control reaction with Pyrosequencing primer; (3) PCR negative control reaction without Pyrosequencing primer; (4) Biotinylated reverse PCR primer alone; (5) Pyrosequencing primer alone; (6) Biotinylated PCR primer with Pyrosequencing primer. Failure to positively genotype samples using SOP[3] designed primers for PCR and Pyrosequencing can often be the result of background signal associated with the negative control reactions. Isolation of the background signal can be helpful in redesigning assays for selection of new primer trio sets.

5. Contact information for web server administration. The SOP[3] application is maintained by the Division of Immunogenetics at the Children's Hospital of Pittsburgh. Inquiries concerning customized searches of the database or reports of errors in the application should be addressed to Steven Ringquist (email: smr73@pitt.edu).

Acknowledgments

This work was supported by grants from the NIH-NLM, U19-AI056374-01 Autoimmunity Centers of Excellence (SR and MT), RO1DK24021(MT) from the National Institutes of Health, and ERHS no. 00021010 (SR and MT) from the Department of Defense. PVB was supported partially by the NSF grant MCB0316255.

References

1. Rozen, S. and Skaletsky, H. (2000) Primer3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols*, (Krawetz, S. and Misener S., eds.), Humana Press, Totowa, NJ, pp. 365–386.
2. Varotto, C., Richly, E., Salamini, F., and Leister, D. (2001) GST-PRIME: a genome-wide primer design software for the generation of gene sequence tags. *Nucl. Acids Res.* **29,** 4373–4377.
3. Emrich, S. J., Lowe, M., and Delcher, A. L. (2003) PROBEmer: a web-based software tool for selection optimal DNA oligos. *Nucl. Acids Res.* **31,** 3746–3750.
4. Stein, L. D. (2003) Integrating biological databases. *Nat. Rev. Genet.* **4,** 337–345.
5. Ringquist, S., Pecoraro, C., Gilchrist, M. S., et al. (2005) SOP[3]v2: web-based selection of oligonucleotide primer trios for genotyping of human and mouse polymorphisms. *Nucl. Acids Res.* **33,** W548–W552.
6. Schildkraut, C. and Lifson, S. (1965) Dependence of the melting temperature of DNA on salt concentration. *Biopolymers* **3,** 195–208.
7. Ronaghi, M., Pettersson, B., Uhlen, M., and Nyren, P. (1998) PCR-introduced loop structure as primer in DNA sequencing. *Biotechniques* **25,** 876–878.
8. Ringquist, S., Alexander, A. M., Rudert, W. A., Styche, A., and Trucco, M. (2002) Pyrosequence based typing of alleles of the HLA-DQB1 gene. *Biotechniques* **33,** 166–175.
9. Utting, M. J., Hampe, J., Platzer, M., and Huse, K. (2004) Locking of 3' ends of single-stranded DNA templates for improved Pyrosequencing performance. *BioTechniques* **37,** 66–73.

Pediatric Diabetes 2007: 8: 307–322

Pediatric Diabetes

Review Article

Navigating pathways affecting type 1 diabetic kidney disease

Pasquali L, Trucco M, Ringquist S. Navigating pathways affecting type 1 diabetic kidney disease.
Pediatric Diabetes 2007: 8: 307–322.

Lorenzo Pasquali[a,b], Massimo Trucco[a] and Steven Ringquist[a]

[a]Division of Immunogenetics, Department of Pediatrics, Rangos Research Center, Children's Hospital of Pittsburgh, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA; and [b]Department of Pediatrics, Institute G Gaslini, University of Genoa, Genoa, Italy

Key words: diabetes – genetics – insulin – nephropathy – renal disease

Corresponding author:
Massimo Trucco, MD
Division of Immunogenetics
Department of Pediatrics
Rangos Research Center, Children's Hospital of Pittsburgh
University of Pittsburgh School of Medicine
3460 Fifth Avenue
Pittsburgh, PA 15213
USA.
Tel: (412) 692-6570;
fax: (412) 692-5809;
e-mail: mnt@pitt.edu

Submitted 4 April 2007. Accepted for publication 10 May 2007

Complications of type 1 diabetes (T1D), even with reduction in risk as a result of better glycemic control, are still the major concern for pediatricians dealing with young patients who during their lifetime will eventually have to face impaired vision, reduced peripheral sensitivity, and kidney disease. T1D is associated with increased risk of nephropathy, a clinical syndrome usually accompanied by other diabetic-related complications such as retinopathy and neuropathy. Blood pressure elevation, cerebrovascular disease and high risk of cardiovascular morbidity and mortality accompany the renal symptoms, completing the picture of the syndrome (1). An estimated 20–40% of T1D patients will develop diabetic nephropathy (DN), clinically first evidenced by microalbuminuria, during their lifetime. If untreated, nearly all T1D patients experiencing microalbuminuria will progress to overt nephropathy, evidenced by macroalbuminuria, followed by declining kidney function, and culminating in end-stage renal disease (ESRD) (2, 3). In addition to frequent morbidity linked to treating complications of the kidney (e.g., dialysis and kidney transplantation), patients experience a high rate of mortality, with a 5-yr survival rate of 21% once ESRD has developed. An estimated 45% of new patients requiring dialysis and kidney transplant are diabetic (T1D and T2D), and diabetes is the most common and rapidly increasing cause of ESRD in the US and European populations.

The prevalence and course of DN are similar in T1D and T2D patients when matched for duration of the disease (4). T2D incidence is increasing in children

and has become a major concern for pediatricians. DN follows an established natural history. The cardinal clinical feature of the syndrome is the progressive increase in urine protein excretion rate (2). The clinical course starting with microalbuminuria through proteinuria and azotemia culminates with ESRD. Before the onset of overt nephropathy, a prolonged period of clinical silence hides various changes in renal function such as hyperfiltration, hyperinfusion, and increasing capillary permeability to macromolecules.

A progressive rise in arterial blood pressure and albuminuria accompanies glomerular filtration rate (GFR) decline. T1D patients predominantly develop diastolic hypertension, whereas T2D patients typically manifest systolic hypertension (1). Microalbuminuria remains the best predictor of DN in both T1D and T2D patients (5). Risk factors correlated with disease progression are poor glycemic control, hypertension, dyslipidemia, elevated serum cholesterol, and smoking (6–8). ESRD is the major cause of mortality in T1D patients and is the dominant indicator of fatality as a result of cardiovascular disease (9, 10).

Changes in kidney structure described in diabetic patients (T1D and T2D) include nodular glomerular sclerosis (Kimmelstiel–Wilson nodular disease) (11) and a diffuse, generalized process of mesangial expansion (diffuse diabetic glomerulosclerosis). In addition, ultrastructural changes such as thickening of the tubular and the glomerular basement membrane (GBM), arteriolar hyalinosis (afferent and efferent) and increase of mesangial matrix component have been observed in both T1D and T2D. Changes in the GBM and the mesangium already can be detected after 2 yr of diabetes during a period of clinical silence (12). However, in some T2D patients, a more heterogeneous renal pathology has been described (13, 14). The latter includes superimposed changes observed in T2DM such as chronic vascular (arteriosclerotic type) and tubulointerstitial lesions or glomerular changes unrelated to diabetes such as proliferative glomerulonephritis or membranous nephropathy (15). The atypical pattern in renal pathology described for T2D may underline the interplay of different factors acting as promoters of progression in the disease other than hyperglycemia.

Genetic risk of DN and ESRD

It is widely accepted that poorly controlled blood glucose while closely correlated with T1DN is insufficient to fully account for the incidence of kidney disease in this population. Approximately half of the patients with poor glycemic control do not develop nephropathy (16–18). Epidemiologic study of the disease has provided evidence supporting the hypothesis that there is a genetic predisposition. Among T1D patients who have healthy kidney function 30 yr after onset of diabetes, the incidence of T1DN decreases, even among individuals experiencing retinopathy. New cases of T1DN are few among individuals with long-lasting T1D (3, 19, 20). Studies of siblings concordant for T1D have demonstrated familial clustering of DN. The increased familial incidence of T1DN allowed the estimation of the genetic risk ratio for a sibling ranging between roughly twofold for T1DN and threefold for T1D–ESRD (21–24).

T1DN is a complex genetic disease indicating that synergy among several genes, with varying individual effects, are contributing to the disease. Models of genetic susceptibility to T1DN are based primarily on the analysis of Rogus et al. (25) who used the data reported by Quinn et al. (23) to simulate the incidence of T1DN among siblings. Their analysis modeled the occurrence of T1DN assuming co-occurrence of 71.5% when the index case exhibited T1DN and 25.4% when the index case presented T1D but with normal kidney function. Comparison of these observations, when overlaid onto a genetic model of an autosomal dominant gene with 20% disease allele frequency and 100% penetrance, implied a 36% lifetime risk of developing nephropathy in the individuals with T1D. Risks of 66.2 and 19.0% were predicted for T1D siblings of the index cases with and without T1DN, respectively. In contrast, an autosomal recessive model assuming 60% disease allele frequency in the Caucasian population also accounted for the observed lifetime risk of T1DN and similar sibling risks of 64.0 and 23.3% for an index case with siblings who are T1DN or T1D, respectively.

Genome-wide mapping aimed at understanding the basis for inherited risk of diseases coincident with T1DN (i.e., T1D and hypertension) have identified genetic elements consistent with their etiology (Table 1). For example, T1D has been strongly linked to inheritance of histocompatibility leukocyte antigen (HLA) class II DRB1 and DQB1 loci expressing a non-aspartate residue at amino acid 57 (26, 27). An estimated 95% of T1D patients have inherited at least one copy of the DRB1*0301,DQB1*0201 or the DRB1*04,DQB1*0302 haplotype (28). Between 30 and 40% of T1D patients are heterozygous for the combined haplotype despite the expected frequency of 2% in their respective populations (29–31). However, HLA is estimated to account for only 40% of the genetic risk of developing T1D (32–34). Genetic analyses of affected individuals have identified additional T1D susceptibility loci, reproducibly implicating chromosomal regions 11p15.5 (insulin-dependent diabetes mellitus [IDDM]2, INS), 2q33 (IDDM12, CTLA4 region), and 1p13 (PTPN22) (34–36).

Comparison of biopsy tissues from the kidneys of T1DN patients supports the hypothesis that familial factors affect severity of the glomerular lesion. In fact

Table 1. Confirmed genetic elements influencing susceptibility to type 1 diabetes

Locus symbol	IDDM region	Cytogenetic band	Odds ratio	Sibling risk ratio	Description
HLA-DQB1	IDDM1	6p21.3	95	2.5–3.6	HLA class II locus DQB1 accounts for greater than 40% of the relative genetic risk associated with T1D
INS	IDDM2	11p15.5	2.7	1.29	Association between a VNTR marker located immediately 5′ of the insulin gene locus
CTLA4	IDDM12	2q33	1.2	1.01	The CTLA4 locus along with ICOS and CD28 are each strong candidates for IDDM12-associated risk
PTPN22	–	1p13	1.8	1.23	The PTPN22 locus encodes a tyrosine phosphatase and is linked with an increased risk for T1D as well as other autoimmune diseases, e.g., rheumatoid arthritis, systemic lupus erythematosus, and thyroiditis

HLA, histocompatibility leukocyte antigen; ICOS, inducible t-cell costimulator; IDDM, insulin dependent diabetes mellitus; INS, insulin; T1D, type 1 diabetes; VNTR, variable number of tandem repeat. Data are summarized from genetic linkage analysis described by Onengut-Gumuscu and Concannon (36) and Motzo et al. (147).

significant similarities between glomerular lesions were identified even when sibling T1D patients were discordant for glycemic control (37). The correlation in severity of glomerular lesions among T1D siblings independent of hemoglobin A1C (HbA1C) levels is consistent with an underlying genetic predisposition. Development of lesions in normal kidneys transplanted into diabetic patients has been described in a long-term study of recipients 6–14 yr after transplantation. Wide variation in the rate of development of lesions in different kidneys was observed independently from the history of blood glucose levels. This suggests that in addition to glycemic control *per se* risk factors intrinsic to the kidney itself are present (38, 39).

In addition to T1D, hypertension is an inheritable risk factor linked to development of kidney disease, indicating that the genetics of T1DN and essential hypertension may be linked (Tables 2 and 3). By controlling blood volume, the kidney is responsible for long-term blood pressure control. Impaired kidney function is a principal component underlying the physiology of essential hypertension (40, 41). Family-based studies have reported an increased incidence of hypertension, 30–40% higher in patients who develop T1DN (42, 43), whereas patients with uncomplicated T1D exhibit relatively lower blood pressures (41). Parents of T1DN patients frequently exhibit increased blood pressure when compared with parents of unaffected T1D patients, suggesting a familial basis linking hypertension with impaired kidney function (43). Studies designed to deconstruct the role of genetics on hypertension have indicated that different

Table 2. Genes linked to Mendelian hypertensive disorders

Locus symbol(s)	Mendelian disorder	Physiological role	Description
HSB11B1	Apparent mineralocorticoid excess	Salt–water homeostasis	11-beta-hydroxysteroid dehydrogenase, type I (autosomal recessive trait)
CYP11B1, CYP11B2	Glucocorticoid-remediable aldosteronism	Salt–water homeostasis	Cytochrome P450, subfamily XIB, polypeptide 1 and cytochrome P450, subfamily XIB, polypeptide 2 (autosomal dominant trait)
SCNN1B, SCNN1G	Liddle syndrome	Salt–water homeostasis	Sodium channel, nonvoltage-gated 1, beta subunit and sodium channel, nonvoltage-gated 1, gamma subunit (autosomal dominant trait)
NR3C2	Mutations in mineralocorticoid receptor	Salt–water homeostasis	Nuclear receptor subfamily 3, group C, member 2 (autosomal dominant trait)
WNK1, WNK4	Pseudohypoaldosteronism type II	Salt–water homeostasis	Protein kinase, lysine-deficient 1 and protein kinase, lysine-deficient 4 (autosomal dominant trait)

Data are summarized from Mullins et al. (45), Lifton et al. (119), and Cowley (148).

Table 3. Candidate loci influencing susceptibility to essential hypertension

Locus symbol	Cytogenetic band	Physiological role	Description
Candidate genes located on chromosome 1q			
AGT	1q42–q43	RAS	Angiotensin I
AVPR1B	1q32	Salt–water homeostasis	Arginine vasopressin receptor 1B
HSD11B1	1q32–q41	Steroid metabolism	11-beta-hydroxysteroid dehydrogenase, type I
NPR1	1q21–q22	Salt–water homeostasis	Natriuretic peptide A type receptor
REN	1q32	RAS	Renin
Candidate genes located on chromosome 3q			
AGTR1	3q21–q25	RAS	Angiotensin II receptor, type 1
CLCN2	3q27–q28	Salt–water homeostasis	Chloride channel 2
DRD3	3q13.3	adrenergic system	Dopamine receptor D3
ECE2	3q28–q29	Endothelin related	Endothelin-converting enzyme 2
KNG1	3q27	Kallikrein–Kinin system	Kininogen 1
MME	3q25.1–q25.2	Salt–water homeostasis	Membrane metallo-endopeptidase (neutral endopeptidase, enkephalinase)

RAS, renin–angiotensin system. Data are summarized from DeWan et al. (44), Hamet et al. (149), and Jeunemaitre et al. (150).

genes may contribute to increased incidence of hypertension in different ethnic groups (44, 45). Genome-wide scans have implicated different regions of chromosome 3 in Caucasian and African-American populations as well as genetic markers on chromosome 1 and 6 among Caucasians (43, 45, 46) (Table 3). Improved understanding of the inherited basis of T1DN may lead to deeper understanding of the molecular mechanisms underlying the disease, leading to improved therapeutic intervention as well as improved understanding of why the majority of T1D patients maintain normal kidney function while a substantial minority go on to exhibit T1DN.

A number of genome-wide linkage scans have been performed in order to evaluate the genetics underlying impaired kidney phenotype. The first genome scans were performed in a Pima Indian cohort with T2D and proteinuria. Genetic analysis of sib pairs identified significant linkage on chromosome 7. Additional loci were identified on chromosomes 3q and 9. Genome scans performed in an African-American cohort affected by T2D and ESRD revealed evidence for susceptibility on chromosomes 3q, 7p, and 18q (Table 4). However, as these studies were performed using DNA obtained from affected sib pairs concordant for both T2D and proteinuria or ESRD, positive signals may have cosegregated with T2D or nephropathy. Following nephropathy only, a study performed using a cohort of 18 large Turkish families with T2D and proteinuria showed strong linkage with chromosome 18q. A recent genome-wide association study has been performed genotyping greater than 80 000 single nucleotide polymorphisms (SNPs) using a cohort of T2DN (case) and T2D (control) patients, identifying genetic association with the engulfment and cell motility 1 gene (*ELMO1*) and a p value of less than 8×10^{-6} on chromosome 7p (47). Osterholm et al. (48) have recently performed a genome-wide linkage analysis using T1D sib pairs discordant for nephropathy. The study identified five genomic regions with possible linkage to T1DN, implicating chromosomes 3q, 4p, 9p, 16q, and 22p. The overlap between chromosome 3q identified from this study and those performed with T2DN cohorts suggests a possible area of focus for future genetic analyses.

The complex etiology of DN reflects the intricate nature of the disease, requiring a combination of alleles of several genes in addition to exposure to environmental factors. The role of shared environmental risk factors has been extensively evaluated, and while their exact nature is unclear, it is likely that similar lifestyles such as eating and smoking habits are more prevalent among siblings compared with unrelated T1D patients. Glycemic control has been proposed to be a risk factor for nephropathy (49). Studies on diabetic twins have indicated that HbA1C levels are at least partially genetically determined, accounting for greater than 60% of the population variance of glycosylated hemoglobin (50). Individuals with good glycemic control, whether through genetics or as a result of rigorous testing and adjustment of insulin doses to control blood glucose levels, are able to reduce their risk of developing diabetic complications, including T1DN as was shown in the Diabetes Control and Complications Trial (DCCT), as explained later in this text.

Risk factors associated with the progression of T1DN

What factors contribute to onset of nephropathy as a complication in T1D patients? Early age of T1D onset may be linked to the increased risk of T1DN (51), but this conclusion has been controversial (52, 53). The observation that the risk for T1DN reaches a peak incidence at ages 25–29 yr and occurs equally in both males and females may indicate that the triggering

Table 4. Candidate loci influencing susceptibility to diabetic nephropathy

Locus symbol	Cytogenetic band	Physiological role	Description
Candidate genes located on chromosome 3q			
NCK1	3q21	Growth factor/cytokine/signal transduction	NCK adaptor protein 1
RAB7	3q21.3	GTP-binding protein related	RAB7, member RAS oncogene family
SLC2A2	3q26.1–q26.3	Carbohydrate transport/metabolism	GLUT2 (solute carrier family 2 (facilitated glucose transporter), member 2
SUCNR1	3q24–q25.1	Metabolism	Succinate receptor 1
Candidate genes located on chromosome 4p			
ADD1	4p16.3	Cytoskeletal component	Adducin 1
Candidate genes located on chromosome 7p			
ELMO1	7p14.2	Apoptosis related	Engulfment and cell motility 1
IGFBP1	7p14–p12	Growth factor/cytokine/signal transduction	Insulin-like growth factor-binding protein 1
IL6	7p21	Growth factor/cytokine/signal transduction	Interleukin 6
NXPH1	7p22	Cell adhesion	Neurexophilin 1
RAC1	7p22	GTP-binding protein related	Ras-related C3 botulinum toxin substrate 1 (rho family, small GTP-binding protein Rac 1)
Candidate genes located on chromosome 16q			
BCAR1	16q23.1	Cell adhesion	Breast cancer anti-estrogen resistance 1
CDH1	16q22.1	Tight junction component	Cadherin 1, type 1, E-cadherin (epithelial)
CDH3	16q22.1	Tight junction component	Cadherin 3, type 1, P-cadherin (placental)
CETP	16q21	Fatty acid/cholesterol transport/metabolism	Cholesteryl ester transfer protein, plasma
CYBA	16q24	Reactive oxygen species metabolism	Cytochrome *b*(-245), alpha subunit
MAPK3	16q11.2	Growth factor/cytokine/signal transduction	Mitogen-activated protein kinase 3
MMP2	16q13	Extracellular matrix component	Matrix metalloproteinase 2
Candidate genes located on chromosome 18q			
BCL2	18q21.3	Apoptosis related	B-cell CLL/lymphoma 2
DSC2	18q12.1	Cell adhesion	Desmocollin 2
SERPINB7	18q21.33	Extracellular matrix component	Serpin peptidase inhibitor, cladeb(ovalbumin), member 7
SMAD7	18q21.1	Growth factor/cytokine/signal transduction	SMAD family member 7

CLL, chronic lymphocytic leukemia; GLUT, glucose transporter; RAB, RAS-associated protein; RAS, renin-angiotensin system; SMAD, homolog of mothers against decapentaplegic. Data are summarized from Shimazaki et al. (47), Osterholm et al. (48), He et al. (143), and Ewens et al. (90).

event(s) may be independent of differences in T1D onset. The triggering event may, however, be linked to puberty and the beginning of adolescence. Puberty-associated deterioration in glycemic control, hormonal changes, as well as psychosocial pressures have come under increased scrutiny for their influence in patients who fail to maintain good blood glucose control (54). Glycemic control often deteriorates during adolescence in individuals with T1D and may result from developmental changes, especially increased growth hormone secretion during puberty, as well as the transition to more autonomous lifestyle typified by adulthood. In some female patients, there is also the emergence of eating disorders and the misuse of insulin as a means to control weight gain (54, 55).

Constitutional factors

Predisposition to the development of DN can be because of genetic susceptibility as well as factors operating *in utero*. Intra-uterine malnutrition has been associated with a reduced number of nephrons (56), increasing the risk of kidney diseases and hypertension in adult life (57). Exposure to unfavorable intra-uterine environment leading to intra-uterine growth retardation (IUGR) was proposed as a risk factor for DN progression. This conclusion follows from the hypothesis that reduced nephron number observed in IUGR patients (58) may lead to reduced functional reserve in those kidneys subjected to potentially toxic agents (e.g., hyperglycemia or smoking). Rossing et al. (59)

observed an increased risk of DN in women with IUGR and T1D. This was confirmed for both sexes in a cohort of Pima Indian type 2 diabetic patients (60). In contrast, the conclusion failed to be reproduced according to observations reported by Eshoj et al. (61) and in a recent study on Finnish T1D patients (62). Contradictory results from different studies have not yet clarified if there is a correlation between IUGR and DN; however, individuals with essential hypertension were shown to have a markedly reduced number of nephrons in addition to an inverse correlation between birth weight and adult systolic blood pressure (63). It is possible that intra-uterine factors may have a direct effect on DN progression later in life although the effect may be indirect, e.g., via hypertension.

Hyperglycemia

Rates of decline in GFR have been linked in T1D patients to increased levels of blood glucose and occur at a rate of roughly 10–14 ml/min/yr in patients with persistent albuminuria (64). This observation has been confirmed in several independent studies, indicating that HbA1C levels are predictive of the risk of subsequent onset of ESRD. In the absence of therapeutic intervention, the majority of patients with sustained microalbuminuria progress to overt nephropathy over a period of 10–15 yr (65–69).

Control of blood glucose levels is critical to the development of the characteristic complications of diabetes, including DN, as has been demonstrated by the outcome of the DCCT. The DCCT was carried out from 1983 to 1993 and assessed the relative effects of intensive and conventional insulin therapy treatments on preventing development and progression of complications (16). Enrolling greater than 1400 adult patients with T1D, the DCCT showed that HbA1C levels could be reduced from 9.1% to 7.2% with a corresponding reduction in progression of renal complications (16). At the end of the study, the incidence of microalbuminuria was reduced by 34% in a cohort in which no evidence of retinopathy or nephropathy was present at baseline. In a second cohort, mild and early complications were present and the DCCT reported a 43% reduction in the incidence of microalbuminuria. The success of intensive insulin therapy for the treatment of diabetes was confirmed by a follow-up study showing persistent long-term reduction in the incidence of T1DN (70).

Compliance with the intensive insulin therapy regime 2 yr after conclusion of the study indicated that 95% of the intensive therapy group reported following intensive therapy regime and 70% of the conventional treatment group now reported following intensive therapy regime. This means that 30% of the conventional treatment patients failed to adopt intensive therapy despite knowledge of the results. Comparison of HbA1C levels after 2 yr post-DCCT indicated that the intensive therapy group experienced HbA1C levels of 7.2% during DCCT and 7.9% 2 yr after conclusion of the trial. Of these patients, 46% continued to monitor blood glucose four or more times a day. Patients from the conventional treatment group who converted to intensive therapy demonstrated HbA1C levels of 8.1% post-DCCT. Of these patients, 36% reported monitoring blood glucose four or more times a day. The benefits of intensive therapy 5 yr post-DCCT were disappointing in that both cohorts reported average HbA1C levels exceeding 8%. One conclusion of the study is that in the absence of a strong research motivation and monthly visits with a health care provider, the degree of intensive treatment shown in the DCCT was not maintained. Moreover, the availability of health insurance corresponded to success as measured by reduction of HbA1C levels. Patients receiving intensive therapy 5 yr post-DCCT without health insurance maintained an average HbA1C of 8.6%, while patients with health insurance maintained an average HbA1C of 8.1%.

Unfortunately, patients participating in intensive insulin therapy treatments for diabetes are more likely to experience severe adverse effects resulting from frequent injections of insulin. Severe hypoglycemic events occurred nearly threefold more often in patients enrolled in the intensive therapy regime and were characterized as requiring assistance of another person. This precluded the enrollment of children into the study. Severe hypoglycemic episodes occurred in 25% of the total number of reported cases of hypoglycemia and were manifested by coma or convulsions. Of the adult patients receiving intensive insulin therapy, 22% experienced at least five episodes of severe hypoglycemia vs. 4% of conventionally treated subjects. Life-threatening complications of hypoglycemia are not uncommon, limiting the availability of intensive insulin therapy as a means to prevent onset of DN.

Hypertension

High blood pressure is prevalent in families in which T1DN is present compared with T1D families without DN (43). Hypertension-associated microalbuminuria is a predictive indicator of overt nephropathy in T1D patients (71). While normalization of blood glucose levels has been linked to healing of microvascular lesions in the kidney, treatment of hypertension has resulted in improved renal outcome by preventing glomerular damage. Clinical studies have indicated a decline in the incidence of nephropathy in T1D patients (72–74). However, many T1D subjects only attain HbA1C levels of 8.5% well in excess of the recommended levels of blood glucose. Regression of microalbuminuria has been observed upon blood pressure control, and remission of overt nephropathy

has been observed in T1D patients treated for hypertension (5, 75).

There is an exponential impact from hypertension on increasing risk for renal disease (76). Maintaining a sustained reduction in blood pressure is the single most influential intervention for reducing progression of nephropathy in T1D as well as T2D patients. Studies of initially normotensive T2D subjects without renal disease who subsequently developed and were treated for hypertension indicated that patients with blood pressure less than 130/80 mm Hg rarely developed nephropathy. In contrast, diabetic patients with blood pressure in the range of 130/80 to 140/90 mm Hg exhibit a significant decline in GFR coincident with 30% of patients progressing to microalbuminuria within 12–15 yr (77). Improvements in blood pressure treatment and glycemic control along with decreased incidence of smoking lead to improved outcomes for T1D patients.

The kidney structures implicated in the pathogenesis of the disease

The glomerular filtration barrier in T1DN

A thorough understanding of the molecular properties of the glomerular filtration barrier (GFB) is crucial to comprehension of the pathology of proteinuria encountered in T1DN. Structurally, the GFB is composed of three layers that mediate the process of ultrafiltration of blood flowing through anastomosing capillary loops leading to the formation of the primary urine in the Bowman's capsule and finally into the tubular system (Fig. 1). The first layer is a highly fenestrated vascular endothelium that covers the wall of the capillaries. It has been suggested that a protein

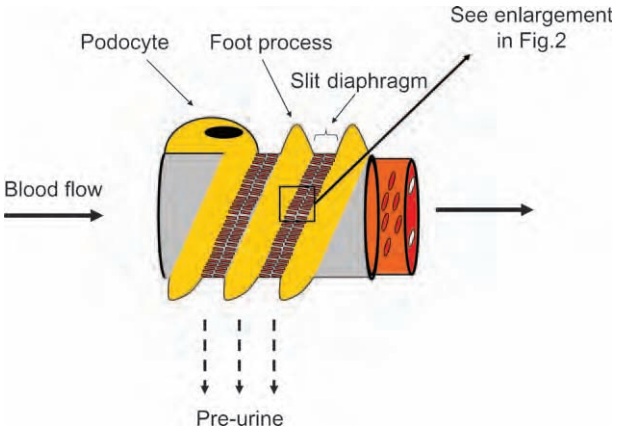

Fig. 1. The GFB. Ultrafiltration occurs when the blood flows through the glomerular capillaries. The GFB comprises the fenestrated endothelial cell layer (red) and the glomerular capillary basement membrane (gray). The podocytes (yellow) loop to the capillaries with their pedicules or foot processes. A protein complex forming the slit diaphragm covers the space between the podocytes processes. GFB, glomerular filtration barrier.

filtration role of the endothelium may be critical to proper kidney function because of the barrier properties of glycocalyx coating of the endothelial cells and filling in of the intracellular space (78). The second layer is the GBM, forming the boundary between blood and urine primarily composed of an interconnecting network of type IV collagen (alpha-1 and -2 during fetal development and alpha-3, -4, and -5 in the adult) and laminin-1 (consisting of alpha-1/beta-1/gamma-1 chains in the fetus) and laminin-11 (consisting of alpha-5/beta-2/gamma-1 chains in the adult). The third layer is composed of highly specialized cells, the podocytes whose cytoplasm extends into interdigitating pedicles or foot processes. The area between the podocyte foot processes is the slit diaphragm composed of a number of cell surface proteins including nephrin, neph 1, and P-cadherin. Understanding GFB function has centered on the role of the slit diaphragm formed by these proteins, serving as the ultimate layer ensuring that large macromolecules such as serum albumin and gamma globulin remain in the bloodstream while allowing small molecules such as water, glucose, and ionic salts to pass through (79).

The genesis of albuminuria is consistent with generalized vascular dysfunction, alterations in extracellular matrix components, and impaired basement membrane organization resulting from hyperglycemia (80). Increasing consideration has been given to the GFB and its central role in the filtration of plasma proteins. High-molecular-weight tracer molecules have been used to investigate the molecular sieving properties of the GFB. Ferritin has an estimated Stokes radius (SR) of 61 Å and when injected into healthy rats failed to pass through the GBM (81). In puromycin aminonucleoside-induced nephrosis models, which induce glomerular injury evidenced by flattening of podocyte foot processes, ferritin is only partially retained (82). Smaller proteins have also been used to examine the selective permeability of the GBM. For example, horseradish peroxidase (SR = 30 Å) has been observed to pass through both the GBM and slit diaphragm, while myeloperoxidase (SR = 36 Å) and catalase (SR = 52 Å) are GBM permeable but are captured by the slit diaphragm prior to entering the urine, consistent with the hypothesis that for large proteins, GBM selectivity is the primary barrier separating the blood and urine space (79).

Genetic polymorphisms of extracellular matrix proteins as well as enzymes involved in their metabolism have been suggested as possible explanations for the variation in susceptibility to kidney disease observed in T1D patients (80, 83). Type IV collagens are a principal component of basement membranes (84). In the adult glomerulus, collagens encoded by the COL4A3, COL4A4, and COL4A5 loci replace the fetal collagens COL4A1 and COL4A2 (85). Absence of the adult forms of collagen has been associated with proteinuria

because of Alport syndrome, a genetically determined disease characterized by heavy proteinuria. In these patients, the pathology is sustained by an abnormal basement membrane because of defects and/or absence of adult type IV collagens (86, 87).

SNPs associated with fetal expressed collagen COL4A1 have also been linked with susceptibility to nephropathy. GBM abnormalities characterized by irregularities in the parietal epithelium lining Bowman's capsule have been observed in heterozygous mice containing a single nucleotide mutation in COL4A1, resulting in a Gly627Trp amino acid substitution in exon 26 (83) as well as mice harboring a G to A polymorphism in a COL4A1-encoded splice site resulting in loss of exon 40 from the mRNA transcript (88). Loss of COL4A1 exon 40 has been linked to nearly 100% prevalence of albuminuria in animals by the time they reach 2 yr of age (89). In humans, mutations affecting GBM structure have been linked with hereditary risk of nephropathy (89, 90). While deletion of COL4A1 is lethal, mice die by embryonic day 9.5 (83), and heterozygous individuals may survive to adulthood carrying increased risk of microvascular complications as a result of altered basement membrane structure occurring during early development and organogenesis.

The third layer and ultimate barrier for plasma protein ultrafiltration is the slit diaphragm of the podocytes

Podocytes, highly specialized epithelial cells, lie on the outer surface of GBM to which they affix by cell surface adhesion proteins such as the alpha-3 beta-1 integrin and dystroglycan (91, 92). The podocyte foot process (polipoid extraflections) cover the outer layer of the GBM. The intrapodocyte connections are typified by narrow spaces (30–40 nm) covered by a porous membrane known as the slit diaphragm (Fig. 2). The slit-diaphragm protein complex is an adherens-type junction and presents a dense zipper-like structure composed of the extracellular components of nephrin, neph 1, P-cadherin and FAT tumor suppressor homolog 1 (FAT 1) (93, 94). Podocin and CD2-associated protein (CD2AP) are localized in the intercellular compartment and interact with the extracellular components of the slit diaphragm anchoring the cytoplasm domain of nephrin to cholesterol-rich regions of the plasma membrane and to the actin filament structure of the podocyte cytoskeleton (79, 95). The protein zona occludens-1 is located at the neph 1 insertion site of the slit diaphragm interacting with the cytoskeleton as well as other components of the cell–cell junction. Recent findings have demonstrated the role of the slit-diaphragm adapter protein NCK (96). NCK binds the cytoplasmatic tail of nephrin and following a phosphorylation signal leads to cytoskeletal reorganization. This finding provides evidence that the central role of the cytoplasmatic region of the nephrin/NCK complex is to act in a signaling transduction capacity, transmitting antiapoptotic signals (via Akt, a serine–threonine kinase) (96), promoting remodeling of the cytoskeletal structure that drive formation of the actin-based foot processes of the podocyte.

The key role of the slit-diaphragm components in maintaining correct permeability function of the GFB is underscored by the results obtained from animal

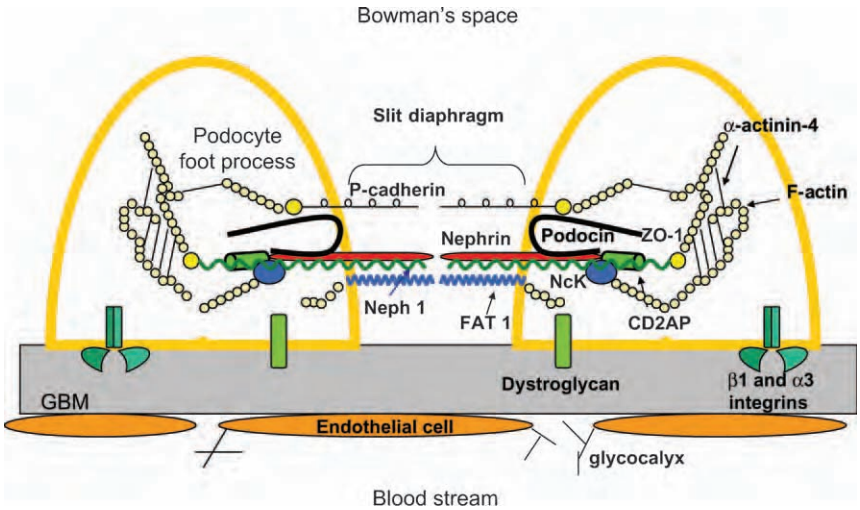

Fig. 2. Schematic view of the molecular anatomy of two podocyte foot processes (yellow) with the interposed slit-diaphragm complex. Interference with proteins that structurally sustain the slit diaphragm (e.g., CD2AP, nephrin, and podocin) can cause foot process effacement resulting in proteinuria. Filtrate from the bloodstream passes through the glycocalyx-filled fenestrae located between endothelial cells lining the GBM. The filtrate then flows through the GBM as well as the slit diaphragm and into the Bowman's space. The image is modified and adapted from Vincenti and Ghiggeri (151) and Johnstone and Holzman (152). CD2AP, CD2-associated protein; FAT 1, FAT tumor suppressor homolog 1; GBM, glomerular basement membrane, NCK, NCK adaptor protein 1; ZO-1, zona occludens-1.

models and human findings. Mutation in genes encoding nephrin, podocin, and α-actinin proteins leads to severe hereditary forms of nephrotic syndrome in humans (97, 98). Mice lacking CD2AP (99) develop nephrotic syndrome, and mutations in CD2AP were found in two human patients with focal segmental glomerulosclerosis. NCK-deficent animals fail in the formation of podocyte foot processes (96).

Progress in elucidating podocyte biology and signaling have shown that the complex of proteins composing the slit diaphragm is critical to maintaining podocyte architecture as well as supporting the overall function of the glomerular filter of the kidney. For T1D, as well as T2D, nephropathy, there are reports of a reduced number and density of the podocytes and an increase in foot process width (100–105). Steffes et al. (102) has reported loss of podocytes occurring early in T1D, while other authors detected the loss of podocytes continuing later in disease progression (104). The loss of podocytes from the glomerular space has been correlated with increased albumin excretion.

A loss of nephrin along with the onset of microalbuminuria occurs in early stages of DN in T1D patients (106). The loss of nephrin was detected even before the appearance of microalbuminuria, suggesting that damage to the slit diaphragm may occur early in disease pathogenesis. Other studies confirmed reduction of nephrin and other slit-diaphragm proteins in the glomeruli of patients with both T1D and T2D nephropathies (107–109). Koop et al. (108) reported on DN in six patients, observing statistically significant decrease in protein level of nephrin along with reduced expression of podocin and podocalyxin. Nephrin mRNA expression inversely correlated with increasing severity of proteinuria (108).

The tubular system

T1DN is often considered a glomerular disease; however, progression to end-stage organ failure also correlates with increased level of renal tubulointerstitial fibrosis (110, 111). Changes occurring in the renal tubular system have been implicated as a primary contributor to the process of renal cortex fibrosis observed in DN (112). Increased fibrosis correlates with macroalbuminuria, nephron tubular cell injury, induction of apoptosis, and hypertension (110, 111, 113–117).

Renal tubular system epithelial cells and interstitial tissue undergo changes during progression of T1DN. These include tubulointerstitium alterations, tubular hypertrophy, and basement membrane alterations that precede tubulointerstitial fibrosis. Tubular cell hypertrophy is an early feature of T1DN. Exposure of proximal tubular epithelial cells to hyperglycemia occurs as a result of glycosuria as well as from high glucose concentration present in interstitial tissue. Cellular response to elevated glucose includes increased tubular basement membrane thickening as a result of increased synthesis of type IV collagen and fibronectin as well as altered matrix protein metabolism as a result of increased expression of metalloproteinase inhibitors tissue inhibitor of metalloproteinase (TIMP)-1 and TIMP-2 (114, 115, 118). Changes to the tubulointerstitium co-occur along with changes to the glomerulus. Poor filtration of the primary urine may lead to presentation of cytotoxic compounds to the nephron tubular cells. Combining the deleterious affects of glycosuria with exposure to cytokines may stimulate cellular responses leading to organ failure.

Primary filtration of urine occurs at the slit diaphragm of the GFB (45); however, glucose and sodium ions are freely filtered. Glucose is entirely reabsorbed in the proximal tubule by a sodium–glucose cotransport system. Recovery of greater than 99% of filtered sodium ion occurs within the nephron via the action of specialized epithelial cells lining the renal tubular system proximal tubule (responsible for 50% sodium absorption), thick ascending loop of Henle (30–40% sodium absorption), distal convoluted tubule (5% sodium absorption), and the collection duct (4% sodium absorption) (Fig. 3). Sodium reuptake by the distal convoluted tubule and collection duct account of a small percentage of total salt reabsorption but are the principal sites at which net salt balance is determined (119). A widely held model for explaining the role of the kidney in blood pressure

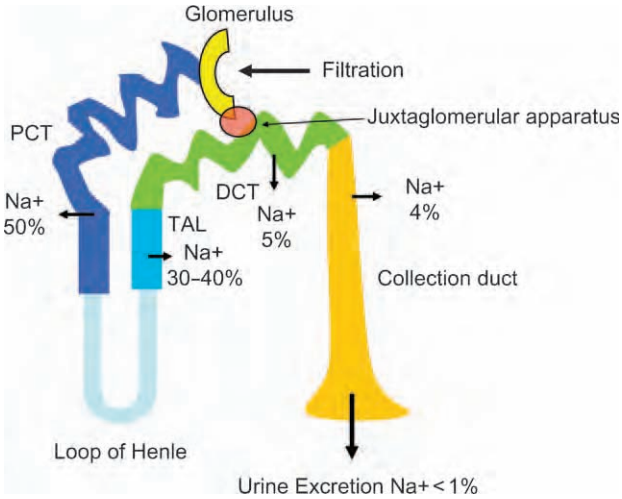

Fig. 3. A schematic view of sodium reabsorption along the nephron. Urine is processed via the PCT, TAL, and DCT. The juxtaglomerular apparatus comprises the macula densa cells, which represent specialized tubular cells at the end of the thick ascending limb, comprising cells from the extraglomerular mesangium, vascular smooth muscle cells, and renin-secreting cells in the media of the afferent glomerular arteriole. DCT, distal convoluting tubule; PCT, proximal convoluting tubule; TAL, thick ascending limb.

control is through its role balancing sodium excretion and recovery. Hypertension results from impaired blood pressure control and correlates with changes in the kidney that lead to impaired renal function, i.e., inability to maintain sodium ion homeostasis.

Increased kidney mass is an early feature of DN (120) with expansion of the proximal tubular cells accounting for most of the growth (121, 122). Along with tubular cell expansion there is increased sodium recovery. Vallon et al. (123) have suggested a model in which increased sodium recovery by the proximal convoluted tubule causes glomerular hyperfiltration, leading to nephropathy. The primary increase in proximal tubular reabsorption in early diabetes may be because of proximal tubular reabsorption through enhanced sodium–glucose cotransport. Studies on patients with T1DN have indicated increased sodium recovery occurring primarily within nephron tubular segments upstream from the macula densa in the distal convoluted tubule (124, 125) and have indicated that there is increased tubular recovery of sodium ions occurring in patients with early-stage T1D.

The juxtaglomerular apparatus comprises different structures in functional and structural link: cells of the extraglomerular mesangium, which fill the angle between the afferent and the efferent glomerular arteriole; vascular smooth muscle cells and renin-secreting cells in the media of the afferent glomerular arteriole; and the macula densa. The macula densa is an area of specialized cells lining the region of the distal convoluted tubule next to the glomerular vascular pole. These cells are sensitive to the ionic content and water volume of the fluid in the distal convoluting tubule. An increase or decrease in sodium, chloride, and potassium uptake elicits inverse changes in the single-nephron GFR by altering the vascular tone, predominantly of the afferent arteriole. The mechanism serves to establish an appropriate balance between GFR and tubular reabsorption upstream from the macula densa. As the proximal tubule grows, more of the glomerular filtrate sodium is reabsorbed and less reaches the macula densa at the end of Henle's loop causing GFR to increase through the normal physiologic actions.

The link between increased tubular sodium recovery and diabetic kidney disease is echoed in observations obtained from studies involving hyperfiltering T1D patients in which there is reported a positive correlation between GFR and sodium recovery in the proximal tubular region of the nephron (125). Examination of T1D patients have indicated that hyperfiltration by the nephron may continue even once blood glucose levels are controlled by intensive insulin therapy (126–128) and may continue to progress independent of glucose control as a result of persisting enlargement of the nephron tubular compartment.

Molecular pathways

Pathways leading to fibrogenesis

The interplay between glomerular hypertension and hyperglycemia may be linked to the pathological changes seen in T1DN. In non-diabetic individuals, glomeruli are protected from deleterious effects associated with essential hypertension by interacting autoregulatory mechanisms (129). Increased blood pressure is associated with increased preglomerular resistance thus isolating glomerular capillaries from potential damage. However, the ability to regulate glomerular blood flow is impaired when significant ablation of renal mass occurs, resulting in increased single-nephron GFR. Carmines et al. (130) demonstrated that regulatory mechanisms are damaged by hyperglycemia-induced inhibition of voltage-gated calcium channels. Endothelial cell injury as a result of hypertensive glomerular damage may be responsible for initiating glomerulosclerosis, with the process involving increased expression of transforming growth factor (TGF)-beta and extracellular matrix proteins (131). The accumulation of excess extracellular matrix within the glomerulus and the interstitium is associated, at least in part, with the profibrotic cytokine, TGF-β (132–134).

The observation that conditions characterized by a severe but relatively selective albuminuria, such as minimal change nephropathy, do not initiate interstitial fibrosis led to the study of the role of higher molecular weight ultrafiltrated proteins in DN. Ultrafiltrated growth factors, such as insulin-like growth factor 1 (IGF-1), hepatocyte growth factor (HGF), and TGF-β, may have a direct effect on the nephron tubular epithelial cells by inducing transcription of genes with profibrogenic effects, whereas other cytokines such as bone morphogenic protein 7 (BMP-7) act as inhibitors of these pathways (135) (Fig. 4). The importance of HGF/IGF-1 action is underscored by the demonstration of a direct effect of these hormones on renal hypertrophy, microalbuminuria, and glomerulosclerosis (136) and that increased excretion rate of both urinary IGF-1 and urinary HGF strongly correlate with microalbuminuria and kidney volume in T1D patients (137).

Pathophysiological mechanisms leading to sclerosis may be related to simultaneous exposure to prosclerotic cytokines in glomeruli and tubulointerstitium as well as tubulotoxicity caused by protein contents in filtrate overloading the proximal tubule's absorption ability. Proteinuria is proposed to be more than a highly reliable predictive factor having a causal role during disease progression (138). Ultrafiltrated proteins undergo endocytosis by the proximal tubular cells, resulting in a cytotoxic effect directly leading to tubular damage. The interaction of filtrated protein

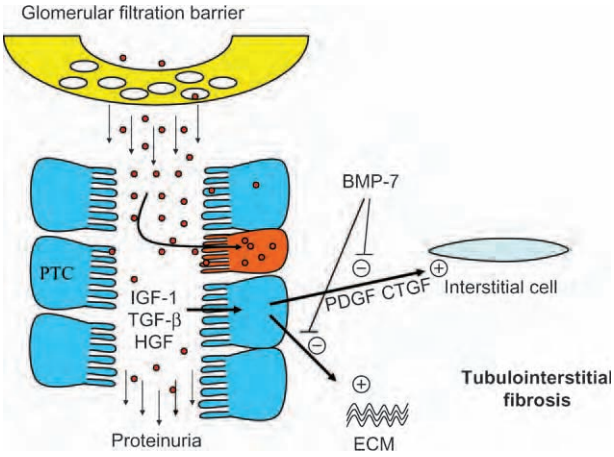

Fig. 4. Interstitial fibrogenesis. When glomerular proteinuria occurs, excess filtrated proteins undergo endocytosis by the PTC (in blue), resulting in cytotoxic effects and tubular damage (red PTC). Ultrafiltrated growth factor cytokines (IGF-1, TGF-β, and HGF) activate a number of genes, enabling PTC to directly contribute to interstitial fibrosis by producing ECM proteins. Tubular cell interaction with growth factor cytokines elicited production of other cytokine signals (PDGF and CTGF) able to induce ECM production by interstitial cells. Other signals such as BMP-7 act as inhibitors of these pathways. BMP-7, bone morphogenic protein 7; CTGF, connective tissue growth factor; ECM, extracellular matrix; HGF, hepatocyte growth factor; IGF-1, insulin-like growth factor 1; PDGF, platelet-derived growth factor; PTC, proximal tubular cells; TGF-β, tissue growth factor-beta.

with proximal tubular cells could be one of the initiating signals enabling a cascade of events leading to changes in extracellular matrix composition and eventually to interstitial fibrogenesis.

Candidate genes and their pathways

Studies of candidate genes for T1DN have focused on genes controlling the balance between extracellular matrix protein synthesis and degradation, glucose metabolism, as well as those whose products control the interaction between kidney function and blood pressure control. Genes influencing the activity of growth factors, protein kinase C activation, hormones regulating blood pressure, accumulation of advanced glycosylation end products, aldose reductase pathway flux, and altered glucose transport have each been extensively evaluated. The roles of hypertension candidate genes have also been studied, although most of the work has focused on the renin–angiotensin system (RAS).

Candidate gene studies have evaluated both functional (i.e., nonsynonymous codons) and allelic markers linked statistically for their influence on T1DN. However, this approach has as yet to convincingly identify risk factors when replicated in independent cohorts (139). The failure to replicate first-stage results from carefully controlled genetic studies possibly reflects the confounding influence of environmental factors on statistical power to detect a true positive signal. Strategies for the selection of candidate genes are improved by consideration of linkage scans along with relevant biologic information about the implicated loci. For example, chromosome 3q has been identified in multiple linkage studies for T1DN (48, 140). This region of the genome has also been identified in studies of T2DN in Pima Indians (141) and in African-American families (142). Other chromosomal regions have also been identified by linkage analysis to T1DN, but chromosome 3q is strongly implicated in hypertension (48) and hypertension is frequently found to co-occur with T1DN. A number of candidate genes are contained within the region of chromosome 3q implicated in these studies (Tables 3 and 4). However, one of the possible candidates, the succinate receptor (SUCNR1), has been linked with carbohydrate catabolism and with control of blood pressure. He et al. (143) reported that the orphan G-protein-coupled receptor encoded by the SUCNR1 locus detected succinate, which is a metabolic intermediate produced by the tricarboxylic acid (TCA) cycle during respiration, and increased succinate levels were causal for increased blood pressure in mice. In diabetic patients, it is possible that local mismatch of energy supply and demand, altered metabolism of TCA-cycle intermediates, or injury may lead to mitochondrial dysfunction and the release of succinate into circulation. Succinate molecules activate receptors in the kidney, causing the release of renin and activation of the RAS. The RAS leads to an increase in blood pressure and altered local blood flow. In healthy individuals, this system might act to regulate local blood flow to match metabolic demands. However, in diabetic patients with impaired metabolic control, it might also result in hypertension or altered cellular function and may contribute to renin-mediated constriction of the renal artery (144).

Many other candidate genes have been examined during genetic evaluation of T1DN. For example, Ewens et al. (90) investigated 110 candidate genes by transmission/disequilibrium testing (TDT) (145) using 72 family trios of which the offspring experienced T1DN. Genes were chosen based on the results of previous studies associating various regions of the human genome to the phenotype of T1DN and have implicated proteins expressed within the GFB. Extracellular matrix proteins encoded by the loci COL4A1, LAMA4, and LAMC1 comprise important components of GBM forming the blood–urine barrier in the kidney. Genetic variants linked to different alleles of the COL4A1 locus have been studied by TDT statistical analysis supporting the hypothesis for genetic linkage of COL4A1 (p value = 0.0002) to the nephropathy phenotype along with additional linkage to the laminin-encoding loci, LAMA4 (p value, 0.016) and LAMC1 (p value, 0.026) (90).

Concluding remarks

The goal of reducing the incidence of T1D complications has been intensively investigated. T1DN is the principal etiology leading to ESRD. The molecular mechanism(s) underlying the disease involve an interplay between genes and gene–environmental exposures. Improved understanding of the disease pathogenesis will lead to better identification of those at risk and by doing so lead to improved treatment regimens.

The evidence for a dominant genetic role in determining susceptibility to kidney disease in T1D patients is primarily the result of epidemiological studies indicating that prevalence of DN increased during the first 15 yr after onset of T1D. After 20 yr duration of diabetes, the incidence of new cases of nephropathy among T1D patients plateaus and in fact may decrease (3, 19, 20). These observations have frequently been interpreted as indicating that there exists a subset of patients susceptible to develop kidney disease. Additional evidence for genetic risk has been obtained from family studies, showing the clustering of DN among T1D siblings (21–24). Siblings experiencing T1D have a significantly increased risk for DN when the T1D proband experiences the disease. Interpretation of the data generated from genetic analysis of T1DN is complicated by the possibility that signals are related to coincident diseases (e.g., hypertension) as well as environmental exposures such as smoking. In order to interpret these results, it is necessary to compare the results of genetic testing of control populations as well as replication of the results in independently recruited case cohorts.

Replication of genome-wide linkage and candidate gene studies for T1DN have been limited by lack of adequate sample size. Recruitment of large independent cohorts for T1DN genetic analysis is required to attain statistical power to detect true genetic signals. Several initiatives are ongoing to address this issue. The most recent of these studies is the genetics of kidneys in diabetes (GoKinD), which has provided a collection of DNA samples and relevant clinical information from greater than 3000 study participants (146). A genome-wide association study of the GoKinD cohort has recently been funded as part of the Genetic Association Information Network's collaborative effort to evaluate genes in complex diseases (http://www.fnih.org/GAIN/GAIN_home.shtml). The initial results of a 500 000 SNP genome-wide association scan based on the GoKinD cohort is due in the fall of 2007. Collection of additional cohorts is ongoing at a number of institutions investigating T1DN genetic risk. Careful collection practices centered on recording relevant health history of participants are essential for success. Knowledge of patient environmental exposures (e.g., history of hypertension, use of anti-hypertensive medications, and smoking) are as essential as recording T1D history and will greatly aid in stratification of populations when comparing participants recruited at different centers.

Acknowledgements

This work was supported by grants U19-AI056374-01 Autoimmunity Centers of Excellence (S. R. and M. T.) and RO1DK24021 (M. T.) from the National Institutes of Health and ERMS #00035010 (S. R. and M. T.) from the Department of Defense.

References

1. Parving HH. Diabetic nephropathy: prevention and treatment. Kidney Int 2001: 60: 2041–2055.
2. Remuzzi G, Benigni A, Remuzzi A. Mechanisms of progression and regression of renal lesions of chronic nephropathies and diabetes. J Clin Invest 2006: 116: 288–296.
3. Pambianco G, Costacou T, Ellis D, Becker DJ, Klein R, Orchard TJ. The 30-year natural history of type 1 diabetes complications: the Pittsburgh Epidemiology of Diabetes Complications Study experience. Diabetes 2006: 55: 1463–1469.
4. Hasslacher C, Ritz E, Wahl P, Michael C. Similar risks of nephropathy in patients with type I or type II diabetes mellitus. Nephrol Dial Transplant 1989: 4: 859–863.
5. Mogensen CE. Microalbuminuria and hypertension with focus on type 1 and type 2 diabetes. J Intern Med 2003: 254: 45–66.
6. Orth SR, Stockmann A, Conradt C et al. Smoking as a risk factor for end-stage renal failure in men with primary renal disease. Kidney Int 1998: 54: 926–931.
7. Chaturvedi N, Fuller JH, Taskinen MR. Differing associations of lipid and lipoprotein disturbances with the macrovascular and microvascular complications of type 1 diabetes. Diabetes Care 2001: 24: 2071–2077
8. Hovind P, Tarnow L, Rossing P et al. Progression of diabetic nephropathy: role of plasma homocysteine and plasminogen activator inhibitor-1. Am J Kidney Dis 2001: 38: 1376–1380.
9. Gall MA, Borch-Johnsen K, Hougaard P, Nielsen FS, Parving HH. Albuminuria and poor glycemic control predict mortality in NIDDM. Diabetes 1995: 44: 1303–1309.
10. Ritz E, Stefanski A. Diabetic nephropathy in type II diabetes. Am J Kidney Dis 1996: 27: 167–194.
11. Kimmelstiel P, Wilson C. Intercapillary lesions in the glomeruli of the kidney. Am J Pathol 1936: 12: 82–97.
12. Drummond K, Mauer M; International Diabetic Nephropathy Study Group. The early natural history of nephropathy in type 1 diabetes: II. Early renal structural changes in type 1 diabetes. Diabetes 2002: 51: 1580–1587.
13. Fioretto P, Mauer M, Brocco E et al. Patterns of renal injury in NIDDM patients with microalbuminuria. Diabetologia 1996: 39: 1569–1576.
14. Schwartz MM, Lewis EJ, Leonard-Martin T, Lewis JB, Batlle D. Renal pathology patterns in type II diabetes mellitus: relationship with retinopathy. The Collaborative Study Group. Nephrol Dial Transplant 1998: 13: 2547–2552.
15. Gambara V, Mecca G, Remuzzi G, Bertani T. Heterogeneous nature of renal lesions in type II diabetes. J Am Soc Nephrol 1993: 3: 1458–1466.

16. DCCT. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. N Engl J Med 1993: 329: 977–986.
17. DCCT. The Diabetes Control and Complications Trial Research Group. Effect of intensive therapy on the development and progression of diabetic nephropathy in the Diabetes Control and Complications Trial. Kidney Int 1995: 47: 1703–1720.
18. Krolewski M, Eggers PW, Warram JH. Magnitude of end-stage renal disease in IDDM: a 35 year follow-up study. Kidney Int 1996: 50: 2041–2046.
19. Krolewski AS, Warram JH, Christlieb AR, Busick EJ, Kahn CR. The changing natural history of nephropathy in type I diabetes. Am J Med 1985: 78: 785–794.
20. Diabetes Research Working Group (DRWG). Conquering diabetes: a strategic plan for the 21st century. NIH Publication No. 99–4398. 1999 (available from http://www.niddk.nih.gov/federal/dwg/fr.pdf). Accessed 6 June 2007.
21. Seaquist ER, Goetz FC, Rich S, Barbosa J. Familial clustering of diabetic kidney disease. Evidence for genetic susceptibility to diabetic nephropathy. N Engl J Med 1989: 320: 1161–1165.
22. Borch-Johnsen K, Norgaard K, Hommel E et al. Is diabetic nephropathy an inherited complication? Kidney Int 1992: 41: 719–722.
23. Quinn M, Angelico MC, Warram JH, Krolewski AS. Familial factors determine the development of diabetic nephropathy in patients with IDDM. Diabetologia 1996: 39: 940–945.
24. Harjutsalo V, Katoh S, Sarti C, Tajima N, Tuomilehto J. Population-based assessment of familial clustering of diabetic nephropathy in type 1 diabetes. Diabetes 2004: 53: 2449–2454.
25. Rogus JJ, Warram JH, Krolewski AS. Genetic studies of late diabetic complications: the overlooked importance of diabetes duration before complication onset. Diabetes 2002: 51: 1655–1662.
26. Todd JA, Bell JI, Mcdevitt HO. HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 1987: 329: 599–604.
27. Morel PA, Dorman JS, Todd JA, Mcdevitt HO, Trucco M. Aspartic acid at position 57 of the HLA-DQ beta chain protects against type I diabetes: a family study. Proc Natl Acad Sci U S A 1988: 85: 8111–8115.
28. Thomson G. HLA DR antigens and susceptibility to insulin-dependent diabetes mellitus. Am J Hum Genet 1984: 36: 1309–1317.
29. Van Der Auwera B, Schuit F, Lyaruu I et al. Genetic susceptibility for insulin-dependent diabetes mellitus in Caucasians revisited: the importance of diabetes registries in disclosing interactions between HLA-DQ- and insulin gene-linked risk. Belgian Diabetes Registry. J Clin Endocrinol Metab 1995: 80: 2567–2573.
30. Rewers M, Bugawan TL, Norris JM et al. Newborn screening for HLA markers associated with IDDM: diabetes autoimmunity study in the young (DAISY). Diabetologia 1996: 39: 807–812.
31. Rich SS, Concannon P. Challenges and strategies for investigating the genetic complexity of common human diseases. Diabetes 2002: 51 (Suppl. 3): S288–S294.
32. Rich SS. Mapping genes in diabetes. Genetic epidemiological perspective. Diabetes 1990: 39: 1315–1319.
33. Ronningen KS, Keiding N, Green A; EURODIAB ACE Study Group. Europe and Diabetes. Correlations between the incidence of childhood-onset type I diabetes in Europe and HLA genotypes. Diabetologia 2001: 44 (Suppl. 3): B51–B59.
34. Concannon P, Erlich HA, Julier C et al. Type 1 diabetes: evidence for susceptibility loci from four genome-wide linkage scans in 1,435 multiplex families. Diabetes 2005: 54: 2995–3001.
35. Concannon P, Gogolin-Ewens KJ, Hinds DA et al. A second-generation screen of the human genome for susceptibility to insulin-dependent diabetes mellitus. Nat Genet 1998: 19: 292–296.
36. Onengut-Gumuscu S, Concannon P. The genetics of type 1 diabetes: lessons learned and future challenges. J Autoimmun 2005: 25 (Suppl.): 34–39.
37. Fioretto P, Steffes MW, Barbosa J, Rich SS, Miller ME, Mauer M. Is diabetic nephropathy inherited? Studies of glomerular structure in type 1 diabetic sibling pairs. Diabetes 1999: 48: 865–869.
38. Mauer SM, Steffes MW, Connett J, Najarian JS, Sutherland DE, Barbosa J. The development of lesions in the glomerular basement membrane and mesangium after transplantation of normal kidneys to diabetic patients. Diabetes 1983: 32: 948–952.
39. Mauer SM, Goetz FC, Mchugh LE et al. Long-term study of normal kidneys transplanted into patients with type I diabetes. Diabetes 1989: 38: 516–523.
40. The Microalbuminuria Collaborative Study Group (MCSG). Predictors of the development of microalbuminuria in patients with Type 1 diabetes mellitus: a seven-year prospective study. The Microalbuminuria Collaborative Study Group. Diabet Med 1999: 16: 918–925.
41. Stern M, Sowers, Tuck ML. Pathophysiology of hypertension in diabetes. In: Le Roith D, Taylor SI, Olefsky JM, eds. Diabetes Mellitus: A Fundamental and Clinical Text, 3rd edn. Philadelphia, PA: Lippincott, Williams, and Wilkins, 2004: 1377–1400.
42. Viberti GC, Keen H, Wiseman MJ. Raised arterial pressure in parents of proteinuric insulin dependent diabetics. Br Med J 1987: 295: 515–517.
43. Krolewski AS, Canessa M, Warram JH et al. Predisposition to hypertension and susceptibility to renal disease in insulin-dependent diabetes mellitus. N Engl J Med 1988: 318: 140–145.
44. Dewan AT, Arnett DK, Atwood LD et al. A genome scan for renal function among hypertensives: the HyperGEN study. Am J Hum Genet 2001: 68: 136–144.
45. Mullins LJ, Bailey MA, Mullins JJ. Hypertension, kidney, and transgenics: a fresh perspective. Physiol Rev 2006: 86: 709–746.
46. Freedman BI, Beck SR, Rich SS et al. A genome-wide scan for urinary albumin excretion in hypertensive families. Hypertension 2003: 42: 291–296.
47. Shimazaki A, Kawamura Y, Kanazawa A et al. Genetic variations in the gene encoding ELMO1 are associated with susceptibility to diabetic nephropathy. Diabetes 2005: 54: 1171–1178.
48. Osterholm AM, He B, Pitkaniemi J et al. Genome-wide scan for type 1 diabetic nephropathy in the Finnish population reveals suggestive linkage to a single locus on chromosome 3q. Kidney Int 2007: 71: 140–145.
49. Alaveras AE, Thomas SM, Sagriotis A, Viberti GC. Promoters of progression of diabetic nephropathy: the relative roles of blood glucose and blood pressure control. Nephrol Dial Transplant 1997: 12 (Suppl. 2): 71–74.
50. Snieder H, Sawtell PA, Ross L, Walker J, Spector TD, Leslie RD. HbA(1c) levels are genetically

determined even in type 1 diabetes: evidence from healthy and diabetic twins. Diabetes 2001: 50: 2858–2863.
51. Harvey JN, Allagoa B. The long-term renal and retinal outcome of childhood-onset type 1 diabetes. Diabet Med 2004: 21: 26–31.
52. Kostraba JN, Dorman JS, Orchard TJ et al. Contribution of diabetes duration before puberty to development of microvascular complications in IDDM subjects. Diabetes Care 1989: 12: 686–693.
53. Svensson M, Nystrom L, Schon S, Dahlquist G. Age at onset of childhood-onset type 1 diabetes and the development of end-stage renal disease: a nationwide population-based study. Diabetes Care 2006: 29: 538–542.
54. Hamilton J, Daneman D. Deteriorating diabetes control during adolescence; physiological or psychosocial? J Pediatr Endocrinol Metab 2002: 15: 115–126.
55. Polonsky WH, Anderson BJ, Lohrer PA, Aponte JE, Jacobson AM, Cole CF. Insulin omission in women with IDDM. Diabetes Care 1994: 17: 1178–1185.
56. Hinchliffe SA, Lynch MR, Sargent PH, Howard CV, Van Velzen D. The effect of intrauterine growth retardation on the development of renal nephrons. Br J Obstet Gynaecol 1992: 99: 296–301.
57. Luyckx VA, Brenner BM. Low birth weight, nephron number, and kidney disease. Kidney Int 2005: 97 (Suppl.): S68–S77.
58. Brenner BM, Chertow GM. Congenital oligonephropathy and the etiology of adult hypertension and progressive renal injury. Am J Kidney Dis 1994: 23: 171–175.
59. Rossing P, Tarnow L, Nielsen FS, Hansen BV, Brenner BM, Parving HH. Low birth weight. A risk factor for development of diabetic nephropathy? Diabetes 1995: 44: 1405–1407.
60. Nelson RG, Morgenstern H, Bennett PH. Birth weight and renal disease in Pima Indians with type 2 diabetes mellitus. Am J Epidemiol 1998: 148: 650–656.
61. Eshoj O, Vaag A, Borch-Johnsen K, Feldt-Rasmussen B, Beck-Nielsen H. Is low birth weight a risk factor for the development of diabetic nephropathy in patients with type 1 diabetes? A population-based case-control study. J Intern Med 2002: 252: 524–528.
62. Fagerudd J, Forsblom C, Pettersson-Fernholm K et al. Low birth weight does not increase the risk of nephropathy in Finnish type 1 diabetic patients. Nephrol Dial Transplant 2006: 21: 2159–2165.
63. Huxley RR, Shiell AW, Law CM. The role of size at birth and postnatal catch-up growth in determining systolic blood pressure: a systematic review of the literature. J Hypertens 2000: 18: 815–831.
64. Parving HH, Rossing P, Hommel E, Smidt UM. Angiotensin-converting enzyme inhibition in diabetic nephropathy: ten years' experience. Am J Kidney Dis 1995: 26: 99–107.
65. Parving HH, Oxenboll B, Svendsen PA, Christiansen JS, Andersen AR. Early detection of patients at risk of developing diabetic nephropathy. A longitudinal study of urinary albumin excretion. Acta Endocrinol (Copenh) 1982: 100: 550–555.
66. Viberti GC, Hill RD, Jarrett RJ, Argyropoulos A, Mahmud U, Keen H. Microalbuminuria as a predictor of clinical nephropathy in insulin-dependent diabetes mellitus. Lancet 1982: 1: 1430–1432.
67. Mathiesen ER, Oxenboll B, Johansen K, Svendsen PA, Deckert T. Incipient nephropathy in type 1 (insulin-dependent) diabetes. Diabetologia 1984: 26: 406–410.
68. Mogensen CE, Christensen CK. Predicting diabetic nephropathy in insulin-dependent patients. N Engl J Med 1984: 311: 89–93.
69. Messent JW, Elliott TG, Hill RD, Jarrett RJ, Keen H, Viberti GC. Prognostic significance of microalbuminuria in insulin-dependent diabetes mellitus: a twenty-three year follow-up study. Kidney Int 1992: 41: 836–839.
70. Writing Team for the Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications Research Group (EDIC). Effect of intensive therapy on the microvascular complications of type 1 diabetes mellitus. JAMA 2002: 287: 2563–2569.
71. Mulec H, Blohme G, Grande B, Bjorck S. The effect of metabolic control on rate of decline in renal function in insulin-dependent diabetes mellitus with overt diabetic nephropathy. Nephrol Dial Transplant 1998: 13: 651–655.
72. Bojestig M, Arnqvist HF, Hermansson G, Karlberg B, Ludvigsson J. Declining incidence of nephropathy in insulin-dependent diabetes mellitus. N Engl J Med 1994: 330: 977–986.
73. Hovind P, Tarnow L, Rossing K et al. Decreasing incidence of severe diabetic microangiopathy in type 1 diabetes. Diabetes Care 2003: 26: 1258–1264.
74. Fagerudd J, Forsblom C, Pettersson-Fernholm K, Groop PH; Finndiane Study Group. Implementation of guidelines for the prevention of diabetic nephropathy. Diabetes Care 2004: 27: 803–804.
75. Hovind P, Rossing P, Tarnow L, Toft H, Parving J, Parving HH. Remission of nephrotic-range albuminuria in type 1 diabetic patients. Diabetes Care 2001: 24: 1972–1977.
76. Giunti S, Barit D, Cooper ME. Mechanisms of diabetic nephropathy: role of hypertension. Hypertension 2006: 48: 519–526.
77. Ravid M, Savin H, Jutrin I, Bental T, Katz B, Lishner M. Long-term stabilizing effect of angiotensin-converting enzyme inhibition on plasma creatinine and on proteinuria in normotensive type II diabetic patients. Ann Intern Med 1993: 118: 577–581.
78. Deen WM, Lazzara MJ, Myers BD. Structural determinants of glomerular permeability. Am J Physiol Renal Physiol 2001: 281: F579–F596.
79. Pavenstadt H, Kriz W, Kretzler M. Cell biology of the glomerular podocyte. Physiol Rev 2003: 83: 253–307.
80. Deckert T, Feldt-Rasmussen B, Borch-Johnsen K, Jensen T, Kofoed-Enevoldsen A. Albuminuria reflects widespread vascular damage. The Steno hypothesis. Diabetologia 1989: 32: 219–226.
81. Farquhar MG, Wissig SL, Palade GE. Glomerular permeability. I. Ferritin transfer across the normal glomerular capillary wall. J Exp Med 1961: 113: 47–66.
82. Farquhar MG, Palade GE. Glomerular permeability. II. Ferritin transfer across the glomerular capillary wall in nephrotic rats. J Exp Med 1961: 114: 699–716.
83. Van Agtmael T, Schlotzer-Schrehardt U, Mckie L et al. Dominant mutations of Col4a1 result in basement membrane defects which lead to anterior segment dysgenesis and glomerulopathy. Hum Mol Genet 2005: 14: 3161–3168.
84. Kalluri R. Basement membranes: structure, assembly and role in tumour angiogenesis. Nat Rev Cancer 2003: 3: 422–433.
85. Miner JH, Sanes JR. Collagen IV alpha 3, alpha 4, and alpha 5 chains in rodent basal laminae: sequence, distribution, association with laminins, and developmental switches. J Cell Biol 1994: 127: 879–891.
86. Hudson BG, Reeders ST, Tryggvason K. Type IV collagen: structure, gene organization, and role in human diseases. Molecular basis of Goodpasture and

Alport syndromes and diffuse leiomyomatosis. J Biol Chem 1993: 268: 26033–26036.
87. Miner JH, Sanes JR. Molecular and functional defects in kidneys of mice lacking collagen alpha 3(IV): implications for alport syndrome. J Cell Biol 1996: 135: 1403–1413.
88 Gould DB, Phalan FC, Breedveld GJ et al. Mutations in Col4a1 cause perinatal cerebral hemorrhage and porencephaly. Science 2005: 308: 1167–1171.
89. Gould DB, Phalan FC, Van Mil SE et al. Role of COL4A1 in small-vessel disease and hemorrhagic stroke. N Engl J Med 2006: 354: 1489–1496.
90. Ewens KG, George RA, Sharma K, Ziyadeh FN, Spielman RS. Assessment of 115 candidate genes for diabetic nephropathy by transmission/disequilibrium test. Diabetes 2005: 54: 3305–3318.
91. Kojima K, Kerjaschki D. Is podocyte shape controlled by the dystroglycan complex? Nephrol Dial Transplant 2002: 17 (Suppl. 9): 23–24.
92. Tsilibary EC. Microvascular basement membranes in diabetes mellitus. J Pathol 2003: 200: 537–546.
93. Ruotsalainen V, Ljungberg P, Wartiovaara J et al. Nephrin is specifically located at the slit diaphragm of glomerular podocytes. Proc Natl Acad Sci U S A 1999: 96: 7962–7967.
94. Reiser J, Kriz W, Kretzler M, Mundel P. The glomerular slit diaphragm is a modified adherens junction. J Am Soc Nephrol 2000: 11: 1–8.
95. Saleem MA, Ni L, Witherden I et al. Co-localization of nephrin, podocin, and the actin cytoskeleton: evidence for a role in podocyte foot process formation. Am J Pathol 2002: 161: 1459–1466.
96. Jones N, Blasutig IM, Eremina V et al. Nck adaptor proteins link nephrin to the actin cytoskeleton of kidney podocytes. Nature 2006: 440: 818–823.
97. Kos CH, Le TC, Sinha S et al. Mice deficient in alpha-actinin-4 have severe glomerular disease. J Clin Invest 2003: 111: 1683–1690.
98. Roselli S, Heidet L, Sich M et al. Early glomerular filtration defect and severe renal disease in podocin-deficient mice. Mol Cell Biol 2004: 24: 550–560.
99. Shih NY, Li J, Karpitskii V et al. Congenital nephrotic syndrome in mice lacking CD2-associated protein. Science 1999: 286: 312–315.
100. Berg UB, Torbjornsdotter TB, Jaremko G, Thalme B. Kidney morphological changes in relation to long-term renal function and metabolic control in adolescents with IDDM. Diabetologia 1998: 41: 1047–1056.
101. Meyer TW, Bennett PH, Nelson RG. Podocyte number predicts long-term urinary albumin excretion in Pima Indians with type II diabetes and microalbuminuria. Diabetologia 1999: 42: 1341–1344.
102. Steffes MW, Schmidt D, Mccrery R, Basgen JM; International Diabetic Nephropathy Study Group. Glomerular cell number in normal subjects and in type 1 diabetic patients. Kidney Int 2001: 59: 2104–2113.
103. Pagtalunan ME, Miller PL, Jumping-Eagle S et al. Podocyte loss and progressive glomerular injury in type II diabetes. J Clin Invest 1997: 99: 342–348.
104. White KE, Bilous RW, Marshall SM et al. Podocyte number in normotensive type 1 diabetic patients with albuminuria. Diabetes 2002: 51: 3083–3089.
105. Dalla Vestra M, Masiero A, Roiter AM, Saller A, Crepaldi G, Fioretto P. Is podocyte injury relevant in diabetic nephropathy? Studies in patients with type 2 diabetes. Diabetes 2003: 52: 1031–1035.
106. Patari A, Forsblom C, Havana M, Taipale H, Groop PH, Holthofer H. Nephrinuria in diabetic nephropathy of type 1 diabetes. Diabetes 2003: 52: 2969–2974.
107. Doublier S, Salvidio G, Lupia E et al. Nephrin expression is reduced in human diabetic nephropathy: evidence for a distinct role for glycated albumin and angiotensin II. Diabetes 2003: 52: 1023–1030.
108. Koop K, Eikmans M, Baelde HJ et al. Expression of podocyte-associated molecules in acquired human kidney diseases. J Am Soc Nephrol 2003: 14: 2063–2071.
109. Benigni A, Gagliardini E, Tomasoni S et al. Selective impairment of gene expression and assembly of nephrin in human diabetic nephropathy. Kidney Int 2004: 65: 2193–2200.
110. Mauer SM, Steffes MW, Ellis EN, Sutherland DE, Brown DM, Goetz FC. Structural-functional relationships in diabetic nephropathy. J Clin Invest 1984: 74: 1143–1155.
111. Bohle A, Wehrmann M, Bogenschutz O, Batz C, Muller CA, Muller GA. The pathogenesis of chronic renal failure in diabetic nephropathy. Investigation of 488 cases of diabetic glomerulosclerosis. Pathol Res Pract 1991: 187: 251–259.
112. Phillips AO. The role of renal proximal tubular cells in diabetic nephropathy. Curr Diab Rep 2003: 3: 491–496.
113. Phillips AO, Steadman R, Morrisey K, Martin J, Eynstone L, Williams JD. Exposure of human renal proximal tubular cells to glucose leads to accumulation of type IV collagen and fibronectin by decreased degradation. Kidney Int 1997: 52: 973–984.
114. Morrisey K, Steadman R, Williams JD, Phillips AO. Renal proximal tubular cell fibronectin accumulation in response to glucose is polyol pathway dependent. Kidney Int 1999: 55: 160–167.
115. Phillips AO, Morrisey K, Steadman R, Williams JD. Decreased degradation of collagen and fibronectin following exposure of proximal cells to glucose. Exp Nephrol 1999: 7: 449–462.
116. Janssen U, Riley SG, Vassiliadou A, Floege J, Phillips AO. Hypertension superimposed on type II diabetes in Goto Kakizaki rats induces progressive nephropathy. Kidney Int 2003: 63: 2162–2170.
117. Erkan E, Garcia CD, Patterson LT et al. Induction of renal tubular cell apoptosis in focal segmental glomerulosclerosis: roles of proteinuria and Fas-dependent pathways. J Am Soc Nephrol 2005: 16: 398–407.
118. Fumo P, Kuncio GS, Ziyadeh FN. PKC and high glucose stimulate collagen alpha 1 (IV) transcriptional activity in a reporter mesangial cell line. Am J Physiol 1994: 267: F632–F638.
119. Lifton RP, Gharavi AG, Geller DS. Molecular mechanisms of human hypertension. Cell 2001: 104: 545–556.
120. Rasch R, Norgaard JO. Renal enlargement: comparative autoradiographic studies of 3H-thymidine uptake in diabetic and uninephrectomized rats. Diabetologia 1983: 25: 280–287.
121. Seyer-Hansen K, Hansen J, Gundersen HJ. Renal hypertrophy in experimental diabetes. A morphometric study. Diabetologia 1980: 18: 501–505.
122. Rasch R, Dorup J. Quantitative morphology of the rat kidney during diabetes mellitus and insulin treatment. Diabetologia 1997: 40: 802–809.
123. Vallon V, Blantz RC, Thomson S. Glomerular hyperfiltration and the salt paradox in early type 1 diabetes mellitus: a tubulo-centric view. J Am Soc Nephrol 2003: 14: 530–537.
124. Brochner-Mortensen J, Stockel M, Sorensen PJ, Nielsen AH, Ditzel J. Proximal glomerulo-tubular balance in patients with type 1 (insulin-dependent) diabetes mellitus. Diabetologia 1984: 27: 189–192.

125. Hannedouche TP, Delgado AG, Gnionsahe DA, Boitard C, Lacour B, Grunfeld JP. Renal hemodynamics and segmental tubular reabsorption in early type 1 diabetes. Kidney Int 1990: 37: 1126–1133.
126. Wiseman MJ, Saunders AJ, Keen H, Viberti G. Effect of blood glucose control on increased glomerular filtration rate and kidney size in insulin-dependent diabetes. N Engl J Med 1985: 312: 617–621.
127. Christensen CK, Christiansen JS, Christensen T, Hermansen K, Mogensen CE. The effect of six months continuous subcutaneous insulin infusion on kidney function and size in insulin-dependent diabetics. Diabet Med 1986: 3: 29–32.
128. Tuttle KR, Bruton JL, Perusek MC, Lancaster JL, Kopp DT, Defronzo RA. Effect of strict glycemic control on renal hemodynamic response to amino acids and renal enlargement in insulin-dependent diabetes mellitus. N Engl J Med 1991: 324: 1626–1632.
129. Keven K, Ekmekci Y. Opposing mechanisms in tubuloglomerular feedback. Am J Kidney Dis 2004: 44: 574.
130. Carmines PK, Ohishi K, Ikenaga H. Functional impairment of renal afferent arteriolar voltage-gated calcium channels in rats with diabetes mellitus. J Clin Invest 1996: 98: 2564–2571.
131. Lee LK, Meyer TW, Pollock AS, Lovett DH. Endothelial cell injury initiates glomerular sclerosis in the rat remnant kidney. J Clin Invest 1995: 96: 953–964.
132. Steffes MW, Bilous RW, Sutherland DE, Mauer SM. Cell and matrix components of the glomerular mesangium in type I diabetes. Diabetes 1992: 41: 679–684.
133. Lane PH, Steffes MW, Fioretto P, Mauer SM. Renal interstitial expansion in insulin-dependent diabetes mellitus. Kidney Int 1993: 43: 661–667.
134. Ziyadeh FN, Sharma K, Ericksen M, Wolf G. Stimulation of collagen gene expression and protein synthesis in murine mesangial cells by high glucose is mediated by autocrine activation of transforming growth factor-beta. J Clin Invest 1994: 93: 536–542.
135. Wang S, Hirschberg R. BMP7 antagonizes TGF-beta -dependent fibrogenesis in mesangial cells. Am J Physiol Renal Physiol 2003: 284: F1006–F1013.
136. Blankestijn PJ, Derkx FH, Birkenhager JC et al. Glomerular hyperfiltration in insulin-dependent diabetes mellitus is correlated with enhanced growth hormone secretion. J Clin Endocrinol Metab 1993: 77: 498–502.
137. Cummings EA, Sochett EB, Dekker MG, Lawson ML, Daneman D. Contribution of growth hormone and IGF-I to early diabetic nephropathy in type 1 diabetes. Diabetes 1998: 47: 1341–1346.
138. Zoja C, Morigi M, Remuzzi G. Proteinuria and phenotypic change of proximal tubular cells. J Am Soc Nephrol 2003: 14: S36–S41.
139. Rich SS. Genetics of diabetes and its complications. J Am Soc Nephrol 2006: 17: 353–360.
140. Moczulski DK, Rogus JJ, Antonellis A, Warram JH, Krolewski AS. Major susceptibility locus for nephropathy in type 1 diabetes on chromosome 3q: results of novel discordant sib-pair analysis. Diabetes 1998: 47: 1164–1169.
141. Imperatore G, Hanson RL, Pettitt DJ, Kobes S, Bennett PH, Knowler WC. Sib-pair linkage analysis for susceptibility genes for microvascular complications among Pima Indians with type 2 diabetes. Pima Diabetes Genes Group. Diabetes 1998: 47: 821–830.
142. Bowden DW, Colicigno CJ, Langefeld CD et al. A genome scan for diabetic nephropathy in African Americans. Kidney Int 2004: 66: 1517–1526.
143. He W, Miao FJ, Lin DC et al. Citric acid cycle intermediates as ligands for orphan G-protein-coupled receptors. Nature 2004: 429: 188–193.
144. Hebert SC. Physiology: orphan detectors of metabolism. Nature 2004: 429: 143–145.
145. Ewens WJ, Spielman RS. The transmission/disequilibrium test: history, subdivision, and admixture. Am J Hum Genet 1995: 57: 455–464.
146. Mueller PW, Rogus JJ, Cleary PA et al. Genetics of Kidneys in Diabetes (GoKinD) Study: a genetics collection available for identifying genetic susceptibility factors for diabetic nephropathy in type 1 diabetes. J Am Soc Nephrol 2006: 17: 1782–1790.
147. Motzo C, Contu D, Cordell HJ et al. Heterogeneity in the magnitude of the insulin gene effect on HLA risk in type 1 diabetes. Diabetes 2004: 53: 3286–3291.
148. Cowley AW Jr. The genetic dissection of essential hypertension. Nat Rev Genet 2006: 7: 829–840.
149. Hamet P, Merlo E, Seda O et al. Quantitative founder-effect analysis of French Canadian families identifies specific loci contributing to metabolic phenotypes of hypertension. Am J Hum Genet 2005: 76: 815–832.
150. Jeunemaitre X, Gimenez-Roqueplo A-P, Disse-Nicodeme S, Corvol P. Molecular basis of human hypertension. In: Rimoin DL, Connor JM, Pyeritz RE, Korf BR, eds. Emery and Rimon's Principles and Practice of Medical Genetics, 5th edn. Philadelphia, PA: Churchill Livingstone, 2007: 1283–1300.
151. Vincenti F, Ghiggeri GM. New insights into the pathogenesis and the therapy of recurrent focal glomerulosclerosis. Am J Transplant 2005: 5: 1179–1185.
152. Johnstone DB, Holzman LB. Clinical impact of research on the podocyte slit diaphragm. Nat Clin Pract Nephrol 2006: 2: 271–282.

On the Use of General Control Samples for Genome-wide Association Studies: Genetic Matching Highlights Causal Variants

Diana Luca,[1,7] Steven Ringquist,[2,7] Lambertus Klei,[3] Ann B. Lee,[1] Christian Gieger,[4,5] H.-Erich Wichmann,[4,5] Stefan Schreiber,[6] Michael Krawczak,[6] Ying Lu,[2] Alexis Styche,[2] Bernie Devlin,[3] Kathryn Roeder,[1,7,*] and Massimo Trucco[2,7]

Resources being amassed for genome-wide association (GWA) studies include "control databases" genotyped with a large-scale SNP array. How to use these databases effectively is an open question. We develop a method to match, by genetic ancestry, controls to affected individuals (cases). The impact of this method, especially for heterogeneous human populations, is to reduce the false-positive rate, inflate other spuriously small p values, and have little impact on the p values associated with true positive loci. Thus, it highlights true positives by downplaying false positives. We perform a GWA by matching Americans with type 1 diabetes (T1D) to controls from Germany. Despite the complex study design, these analyses identify numerous loci known to confer risk for T1D.

Introduction

Systematic GWA studies are critically dependent on the availability of very large and well-characterized control populations. With a different degree of structure in modern populations, ideally, multiple, diverse, and large control populations will be used. As platforms for GWA become standardized, numerous sources of pregenotyped control individuals are becoming available. Typically, many more controls are available than cases, and we believe these controls can be advantageous for discovering risk loci and for controlling the false-positive rate. For example, the data analyzed here include 416 Americans of European descent diagnosed with T1D (MIM 222100) and a control database of 2159 individuals from different regions of Germany.

Ancestry matching based on nongenetic variables[1] and SNP genotypes[2] for genetic-association studies has been proposed previously. Our approach, which we call genetic matching or GEM, goes further in that we show how to systematically obtain favorable matching by using a panel of genetic markers and how to determine outlying individuals as well as individuals that cannot be successfully matched to others in the available registry. By simulations, we will contrast matching to a commonly used method for controlling the confounding of ancestry, namely the use of eigenvector analysis[3] via Eigenstrat[4] to identify predictors of ancestry; for the real data, we contrast matching to both Eigenstrat and identification of common ancestry, such as European American.

We propose matching on the basis of genetic similarities derived from eigenvector decomposition (EVD), making our initial analyses similar to that taken in Eigenstrat.[4] The best known form of matching is matched pairs (pMatch); however, assuming the criterion for matching are sufficient to remove the effects of unmeasured confounding, an alternative to matched pairs known as full matching (fMatch) is optimal.[5] Consider a scenario in which three cases (a, b, and c) and three controls (x, y, and z) fall into two distinct ancestral clusters (a, x, and y) and (b, c, and z). Matching pairs creates three strata, (a and x), (c and z), and (b and y), but the pair (b and y) does not define a homogeneous strata. Alternatively, fMatch minimizes the total distance between individuals within strata with the constraint being that each stratum includes a single case and one or more controls, or vice versa, i.e., clusters (a, x, and y) and (b, c, and z). Of the two, fMatch is optimal because case and control samples are unlikely to have identical distributions of ancestry, and in this situation, forcing each case to match a unique control leads to suboptimal matches. (pMatch can be very useful, however, in designing follow-up studies that require preselection of case and control samples.)

In large association studies, the sample typically includes some individuals with widely varying ancestry. EVD is highly sensitive to outlying observations. A few points lying far from the majority of the data can determine multiple principal axes of the representation. Indeed, outliers can obscure the discovery of axes that potentially separate the data into distinct types. For this reason, individuals

[1]Department of Statistics, Carnegie Mellon University; Pittsburgh, PA 15213, USA; [2]Division of Immunogenetics, Department of Pediatrics, Children's Hospital, Pittsburgh, PA 15213, USA; [3]Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213, USA; [4]GSF National Research Center for Environment and Health, Institute of Epidemiology, Neuherberg 85764, Germany; [5]Ludwig-Maximilians-University Munich, IBE, Chair of Epidemiology, Munich 85764, Germany; [6]Institute of Clinical Molecular Biology and PopGen biobank, Christian Albrechts University, Kiel 24105, Germany
[7]These authors contributed equally to this work.
*Correspondence: roeder@stat.cmu.edu
DOI 10.1016/j.ajhg.2007.11.003.

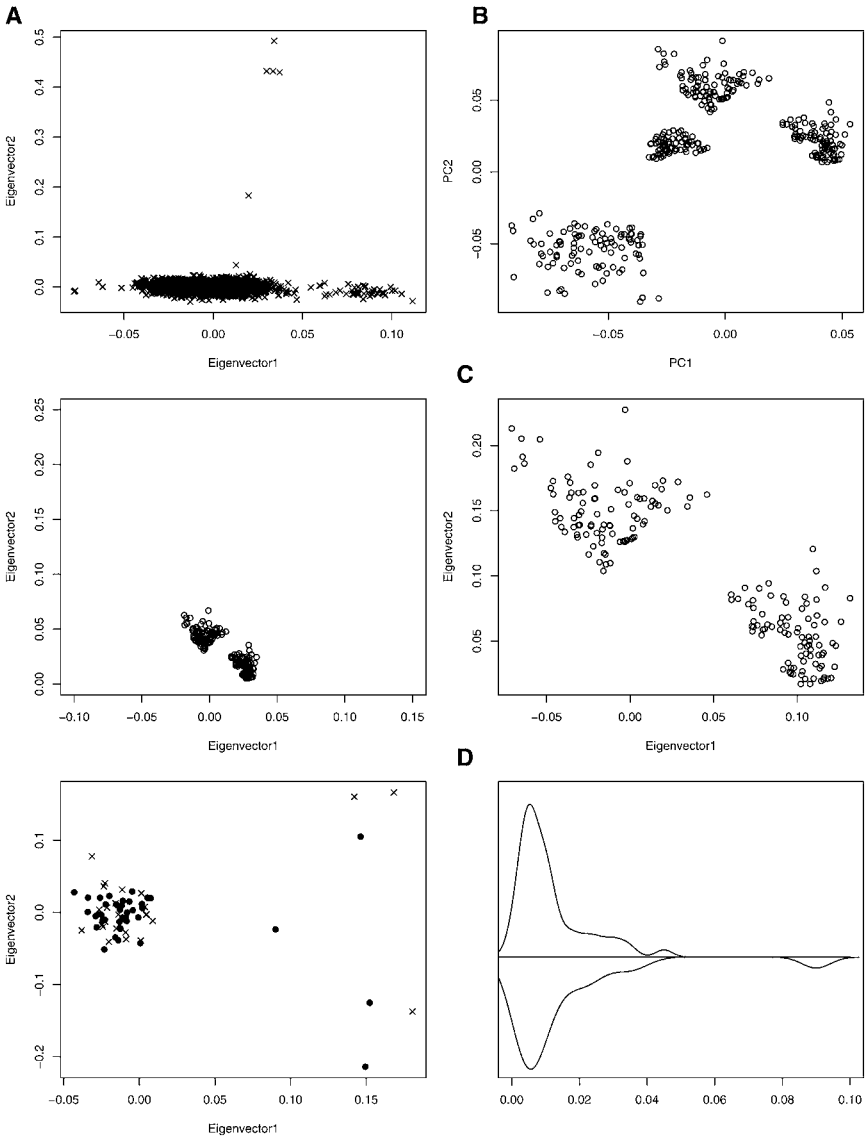

Figure 1. Flowchart for Genetic-Matching Algorithm Illustrated with Portions of the T1D Data

Distances between individuals are determined by the major axes of variation in the EVD representation. Outlier removal, illustrated by (A), is critical for revealing the subtle variability between individuals of similar ancestry. After major outliers are removed, clustering is used for discovery of homogeneous clusters; four distinct clusters are displayed here (B), plotted as principal component axes. Two of these clusters are displayed before ([C], left) and after ([C], right) rescaling of axes. Some observations are not outliers, but they are unmatchable ([D], left); for example, the isolated case in the center of the plot. Rescaled distances are compared to distances expected in homogeneous samples ([D], right) to identify cases and controls that can not be successfully matched. Association analysis is performed on matched strata so that the effects of population structure could be removed (not shown).

having highly unusual measures on any of the major eigenvectors are removed.[4] Likewise, with matching it is necessary to determine which strata span an unusual distance leading to "unmatchable individuals." If the controls are more numerous than the cases, they typically span a larger range of ancestries than cases, and it should be possible to find one or more controls similar to each case. Conversely, some cases may have to be removed to account for the effects of structure. In this work, we formalize the notions of outlying and unmatchable individuals and propose a method to discover the key axes that describe the population structure.

Material and Methods

A Sketch of the Matching Procedure Employed by GEM and Displayed in Figure 1

The illustration (Figure 1) shows the steps involved in matching genotyped cases and controls. To begin, create an L SNPs and N individuals matrix of scaled allele counts from which the EVD is computed (see Appendix). The top D eigenvectors form a D dimensional map describing the "ancestry" of each individual, i.e., the mapping of the ith subject in each dimension is determined by the ith element the in dth eigenvector. The dth eigenvalue determines the importance of the dth dimension in the new representation of the data. Individuals of similar ancestry map to similar values in the eigenvectors associated with large eigenvalues (Figure 1). Eigenvectors associated with small eigenvalues have little or no genetic interpretation.

For a homogeneous population, the largest eigenvalues provides the basis for a significance test for population structure (see Patterson[6] and Appendix). Applying this test with significance level $\alpha = 0.01$, we determine the number of dimensions D to be used in the eigenvector representation. The EVD determines the distance between individuals on the basis of the top D eigenvectors, serving as coordinates or dimensions, and eigenvalues serving as weights to exaggerate differences in dimensions of greater importance (see Appendix).

If the data have many outliers, D will be relatively large, and the principal eigenvectors will be poorly estimated.[7] Outliers (Figure 1A) can be removed with visual diagnostics or the criterion from Eigenstrat;[4] namely, remove any individuals with ancestry coefficients greater than 6 SDs in at least one of the D eigenvector axes. After removing outliers, the EVD should be recomputed. If the estimated dimension, D, is still greater than two or three, we suggest finding the distance between nearest pairs of controls and cases. A subject with ancestry that does not lend itself to matching will appear as an outlier via this criterion and should be removed (see the T1D example).

To determine how to match and which individuals are unmatchable, we rely on the distribution of distances between individuals

in a homogeneous population. For a homogeneous sample, the distribution of distances will depend on sample size N and the number of loci L. By using simulations, we can find the distribution of distances for a homogeneous population. These simulations also yield the distribution of eigenvalues for a homogeneous sample of size N.

Real populations are heterogeneous but can be modeled as mixtures of relatively homogeneous subpopulations (Figure 1B). We wish to represent these subpopulations so that the between-subject distances within a homogeneous subpopulation are comparable to expectation if the entire sample were homogeneous. To do so, we need to model the underlying population substructure and adjust real data so that they are scaled properly (Figure 1C); otherwise, the between subpopulation variance will cause distances between individuals to be poorly calibrated (Figure 1C). We do this via a two-stage algorithm involving clustering and scaling. In stage one, we cluster individuals that appear to have common ancestry. This is done iteratively, by addition of clusters and then testing for structure (see Appendix for testing) until each cluster is homogeneous. We use Ward's algorithm[8,9] to form hierarchical groups of mutually exclusive subsets based on the first D axes of the EVD. We need a stopping rule for choosing K, the number of clusters. Start with $K = 2$ and apply the test for population structure on each of the clusters ($\alpha = 0.001$). Homogeneous clusters, as judged by the significance test, are set aside, and Ward's algorithm is applied only to the remaining data. Repeat this process, increasing K until all the clusters are homogeneous or consisting of too few observations (~20). Finally, we rescale interindividual distances as described in the Appendix so that they are comparable to distances found in a homogeneous population. At this rescaling step, unmatchable individuals are uncovered and removed (Figure 1D).

After outliers and unmatchable individuals are removed from the sample, recalculate the EVD and determine D. Reverting back to unscaled eigenvectors, find the distance between cases and controls on the basis of the Euclidean distance with D dimensions as described in the Appendix. Match strata with either full match or pair match. Software implementing matching algorithms is widely available (e.g., we use the optmatch function in the statistical package R). Then, the data can be analyzed for disease and SNP association by conditional logistic regression. Other covariates can be entered into the model at this point.

T1D Analyses

Purified samples of genomic DNA were obtained from the Genetics of Kidneys in Diabetes (GoKinD) study[10] and from T1D patients recruited at the Children's Hospital of Pittsburgh (CHP) and University of Pittsburgh Medical Center. The study employed a human gene-chip microarray (Affymetrix, Santa Clara, CA) for evaluation of genetic variants with DNA samples from T1D (case) participants with genetic typing data obtained from the KORA[11] and PopGen[12] "control" cohorts.[13] Genotyping results were obtained with the same Affymetrix 500K SNP typing array; however, assays for case and control cohorts were performed independently. Case participants (n = 416) were recruited in the U.S., with self-declared European ancestry and T1D; control participants (n = 2159) were citizens of Germany recruited independent of phenotype (Table 1). Recruitment of participants at CHP was governed by the human subjects protocol approved by the University of Pittsburgh Institutional Review Board (IRB #011052: New Advanced Technology to Improve Prediction of Type 1 Diabetes). CHP patients (n = 28) consented to providing 10 ml blood obtained by vein puncture as well as a brief medical history relating to onset of T1D. The GoKinD cohort (n = 394) was recruited independently from the CHP cohort by collaborative efforts of the Juvenile Diabetes Research Foundation, National Institutes of Health, and U.S. Center of Disease Control.[10] Material from the GoKinD cohort was provided as solutions of DNA, purified from lymphoblastoid cell lines or from whole blood. DNA solutions were provided as 50 μl aliquots containing ~100 ng/μl DNA per aliquot dissolved in 20 mM NaCl and 1 mM EDTA (pH 7.5). DNA from the CHP samples were obtained from whole blood with methods described in Ringquist,[14] and genotyping was performed by Affymetrix Services Laboratory (Affymetrix) with GeneChip 500K arrays. All of the genotype data from GoKinD samples generated by this project will be submitted to an accessible database, such as dbGaP or T1Dbase (see Web Resources).

Table 1. Characteristics of Case and Control Participants

	Case Participants		Control Participants	
Demographic Characteristics	CHP	GoKinD	KORA	POPGEN
Number of singletons	28	394	1644	500
Nominal European American (%)	100%	100%	—	—
German residents (%)	—	—	100%	100%
Male gender (%)	50%	46.7%	49.5%	51.8%
History of Diabetes				
Type 1 diabetes (%)	100%	100%	—	—
Mean age at T1D Diagnosis (yr)	12.7 ± 7.9	12.2 ± 7.1	—	—

All T1D samples had a sufficient completion rate (>95%) for inclusion, as did almost all KORA and PopGen samples. Initially, genotypes for all three samples were called with the BRLMM algorithm.[15] By using three criteria for genotype QC per SNP—greater than 90% genotype calls, test statistic for Hardy-Weinberg yields p value > 0.005, and minor allele frequency ≥ 0.05—we removed ~140,000 SNPs and retained 360,000 for the T1D sample, similar to other studies. When we contrasted the T1D samples to the control samples, we noted SNPs with very different allele frequencies that were not in or near known T1D loci. Inspection of the allele frequencies showed that the control allele frequencies were remarkably similar to HapMap frequencies (see Web Resources), but the corresponding genotype clusters for the T1D samples had undesirable features.

We tried various ways to improve the genotype calls. First, we looked for substantial differences between the calls by using the two algorithms employed by Affymetrix, namely DM and BRLMM. Although some discrepancies were noted, we did not see a material improvement in the data by eliminating this small set of loci. Next, because we had the Affymetrix "cel" files for the PopGen control sample, we called all of these genotypes for PopGen and T1D together by using both the DM and BRLMM algorithms. Again, this process eliminated some problematic loci, but the results were not compelling. Finally, we tried the new Bayesian calling algorithm, CHIAMO.[16] This algorithm led to a marked improvement for the genotype calls, as determined by inspection of the genotype clusters. For our data, we found that analyzing the PopGen and T1D data together (batch) yielded slightly better results than analyzing the two data sets as complementary strata, so these were the data we reported. Because we had greater confidence in the BRLMM calls for chromosome X, we reported those calls for X-linked SNPs. Because the KORA sample came to us only with

genotypes called by the BRLMM algorithm, we used those genotypes for that data set.

Preliminary quality control consisted of a six-step process that reduced the number of cases to 415, controls to 2112, and SNPs to 284,216. Step 1: Removed a case who was a clear outlier. Step 2: Removed 32 controls who had greater than 5% missing genotypes. Step 3: Removed 90,732 SNPs with >5% noncall rate in at least one of the three samples. Step 4: Removed 105,658 SNPs with minor allele frequencies less than 0.05 in either control sample. Step 5: Removed 1972 SNPs with $F_{ST} > 0.02$ (estimated for the two German control samples). Step 6: Removed 18,427 SNPs that violated Hardy-Weinberg equilibrium ($p < 0.005$) in either of the control samples.

Results

Simulations

We compare three approaches to correct for the effects of structure: Eigenstrat and GEM with fMatch and pMatch. Although we compare their size (i.e., rate of false positives) and power, these approaches are not direct competitors. The GEM methods are designed to limit analysis to strata that are chosen a priori, whereas Eigenstrat aims to remove the effects of structure in the analysis stage.

Allele frequencies for the subpopulations were generated with the "Balding-Nichols" model[17] (see Appendix), with allele frequencies varying uniformly between 0.05 and 0.5. To correct for structure, L reference SNPs were generated. Of these SNPs, 99% had a minor amount of variability across subpopulations ($F_{ST} = 0.01$), and 1% had substantial differentiation ($F_{ST} = 0.1$). Null or causal candidate SNPs of three levels of F_{ST} were generated: Model (1) strongly differentiated SNPs, $F_{ST} = 0.1$; Model (2) moderately differentiated SNPs with $F_{ST} = 0.03$; and Model (3), modestly differentiated SNPs with $F_{ST} = 0.01$.

Ten panels of independent reference SNPs, with L ranging from 96 to 100,000, were generated. For each of these panels, we simulated 1000 independent causal SNPs and 1000 independent null candidate SNPs. We repeated this analysis for models (1), (2) and (3) and for six choices of L. Causal SNPs with relative risk $R = 2$ were generated with the approach described in Price[4] for power calculations.

Our first battery of simulations is based on SNPs sampled from two subpopulations, with 200 individuals per subpopulation. Case status was assigned to 80 and 20 of the individuals from subpopulations 1 and 2, respectively. The remaining individuals were assigned control status. For the matched-pairs analysis, we paired each case to the closest control until we obtained 100 matched pairs. For the other two methods of analyses, we analyzed all 400 individuals. Each method readily detects population substructure and achieves the desired type I error rate as L increases (Table 2). pMatch and fMatch successfully remove the effect of structure with a smaller panel of reference SNPs than Eigenstrat does (Figure 2A). Indeed, when a large panel of reference SNPs is available, the GEM procedures are overly conservative; consequently, Eigenstrat is slightly more powerful than both matching procedures (Table 2) under these conditions. For SNPs with less information about population membership than present in our simulated reference panels, greater numbers of SNPs would be required to remove the effects of structure.[4]

Our second battery of simulations is based on nine subpopulations distributed along a gradient, designed to simulate a cline such as the north to south cline observed in western Europe. The 100 cases are distributed with 2, 4, 6, 7, 9, 12, 15, 20, and 25 individuals in populations 1–9, respectively. The 300 controls are distributed randomly across the nine subpopulations. Results from this simulation are qualitatively similar to those shown in Figure 2A (Table 2). The first two batteries of simulations illustrate that when the case and control samples are drawn from the same subpopulations, but with different frequencies, the effects of substructure can be removed with any of the three methods described. Even the effects of highly differentiated SNPs can be removed provided the reference panel is sufficiently informative.

Our third battery of simulations is also based on a nine population gradient; however, the cases and controls are apportioned in a manner that simulates the complexity of human populations and GWA designs. As in the previous simulation, we simulate nine populations and draw 300 controls randomly. In contrast, all 50 of the cases are drawn from populations 6–9. Because of the nature of this third battery, namely the presence of unmatchable observations, we analyze the data in two ways: Unmatchable observations are removed as described previously; or unmatchable observations are retained. In choosing only a single control for each case, pMatch includes only 50 of the controls in the study regardless of the treatment of outliers. Provided the reference panel is sufficiently informative, many of these controls will be derived from populations 6–9. Eigenstrat, on the other hand, uses all of the data, as will fMatch when unmatchable observations are retained. For fMatch, this means that cases drawn from population 6 will tend to have many controls in their strata sampled from populations 1–5. The remaining cases will tend to have only one or two controls in their strata. By grouping the outlying observations, fMatch attempts to minimize the effect of unmatchable observations. Eigenstrat must account for controls sampled from populations 1–5 with regression techniques, which are well known to suffer adverse consequences when they are extrapolating beyond the range of the data.

When unmatchable observations are retained, pMatch corrects for the effects of substructure with fewer reference SNPs than the other two methods (Table 3 and Figure 2B). Indeed, Eigenstrat fails to remove the effects of population substructure. By comparing pMatch and fMatch, we see that the latter has greater power. This makes sense because fMatch is using more of the data (Table 3).

On the basis of the clustering and rescaling process, most of the controls from populations 1–5 are unmatchable, and

Table 2. Size and Power of Tests at Level 0.05

Statistic	Design	No. of Markers	Eigenstrat with F_{ST}			pMatch with F_{ST}			fMatch with F_{ST}		
			0.01	0.03	0.1	0.01	0.03	0.1	0.01	0.03	0.1
Size											
	Two Populations										
		96	.069	.106	.211	.062	.100	.202	.065	.101	.206
		386	.055	.061	.085	.044	.045	.051	.047	.049	.054
		1536	.052	.054	.055	.046	.045	.045	.047	.047	.046
		6144	.053	.052	.051	.044	.045	.045	.047	.047	.046
		12000	.053	.052	.052	.044	.044	.045	.047	.045	.047
		24000	.053	.050	.051	.043	.042	.041	.046	.046	.045
	Gradient										
		96	.069	.109	.221	.049	.067	.109	.061	.097	.201
		386	.054	.058	.071	.043	.046	.048	.048	.052	.063
		1536	.052	.051	.050	.045	.044	.044	.047	.047	.046
		6144	.052	.051	.050	.045	.045	.044	.046	.045	.047
		12000	.052	.052	.052	.044	.044	.045	.047	.047	.047
		24000	.052	.052	.052	.044	.045	.044	.047	.046	.046
Power											
	Two Populations										
		96	.783	.710	.683	.693	.635	.620	.754	.685	.659
		386	.767	.701	.682	.682	.632	.621	.735	.673	.653
		1536	.766	.702	.677	.683	.635	.622	.736	.674	.653
		6144	.765	.694	.676	.684	.630	.623	.735	.671	.653
		12000	.765	.697	.676	.684	.633	.624	.735	.673	.653
		24000	.763	.696	.677	.684	.632	.624	.734	.671	.653
	Gradient										
		96	.939	.917	.833	.886	.872	.804	.922	.900	.814
		386	.924	.891	.796	.877	.857	.782	.902	.869	.775
		1536	.917	.876	.774	.876	.850	.775	.894	.856	.754
		6144	.913	.874	.768	.873	.849	.773	.892	.853	.747
		12000	.915	.873	.771	.874	.849	.774	.891	.850	.749
		24000	.912	.874	.768	.873	.849	.771	.892	.852	.747

Columns depict the results as F_{ST} varies (0.01, 0.03, and 0.1) in the candidate markers. Results are shown for two scenarios: a two-population mixture and a nine-population gradient. For the size, the expected number of p values smaller than 0.05 is 50.

such a result is desirable because cases were only drawn from populations 6–9. In this instance, the size of the matched analyses is now closer to the nominal level even when L is small, as expected. Interestingly, there is the considerable enhancement in power for fMatch and pMatch when unmatchable individuals are removed, as recommended by our methods, as opposed to when they are forced to be retained (Table 3). This occurs because removal of the outliers leads to improved performance of the EVD and hence superior choices of matches in the analysis. In addition, for fMatch the removal of controls from populations 1–5 leads to a more homogeneous sample that tends to increase power.

Eigenstrat defines outliers without specific reference to cases and controls; thus, none of the observations are unmatchable observations. Nevertheless, if the regression approach is applied after removal of those observations declared unmatchable by the fMatch procedure, the type I error is successfully controlled, and the power is slightly greater than it is for fMatch (Table 3). This hybrid approach to analysis has some potential for further development.

GWA of Type 1 Diabetes Data with fMatch

We analyzed 416 cases of T1D,[18] derived from the GoKinD[10] cohort (n = 394) and T1D patients recruited from the Children's Hospital of Pittsburgh (n = 28). Samples were genotyped with the Affymetrix 500K GeneChip. All identified their ancestors as European. The mean age of onset for T1D was 12.2 and 12.7 years of age for the GoKinD and Pittsburgh cohorts, respectively. Controls genotyped by the same chip were obtained from the PopGen and KORA repositories, which consist of 500 individuals from north Germany (PopGen) and 1644 individuals from southern Germany (KORA).[11–13,19] The four cohorts were recruited independently of one another. The relevant characteristics of these cohorts are summarized in Table 1.

Stringent quality control reduced the number of SNPs to 284,216 and the number of controls to 2112 (samples were removed if the rate of missing genotypes exceeded 5%). To reconstruct ancestry, we chose 23,552 independent or "tag" SNPs by using the H-clust algorithm[20] with an r^2 cutoff value of 0.04. Both case and control individuals exhibit complex population heterogeneity. For example, individuals were included in the PopGen and KORA registry on

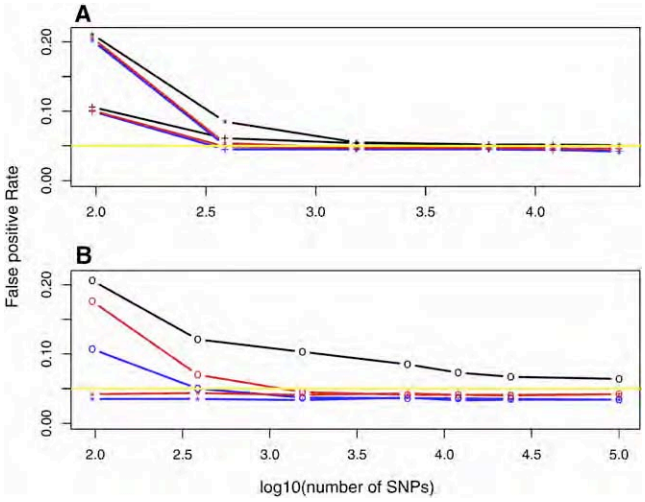

Figure 2. False-Positive Rate versus Log of the Number of Markers Available for Estimating Structure

Results are for Eigenstrat (black), pMatch (blue), and fMatch (red). The desired nominal rate of 0.05 is plotted as a yellow line. In (A), a sample derived from two simulated populations is shown. Results are displayed for markers with two levels of differentiation $F_{ST} = 0.1$(*) and 0.03 (+). The former exhibits a higher rate of errors than the latter for small numbers of markers. In (B), a sample derived from a gradient of simulated populations is shown. Results are displayed for the full sample (plotting character "o") and with unmatchable individuals removed (plotting character "*"; this applies to the matching methods only).

Table 3. Size and Power of the Tests before and after Removing Outliers, with Eigenstrat, pMatch, and fMatch

	Outliers Present								
	Eigenstrat with F_{ST}			pMatch with F_{ST}			fMatch with F_{ST}		
Statistic No. of Markers	.01	.03	.1	.01	.03	.1	.01	.03	.1
Size									
96	.064	.097	.206	.044	.056	.107	.056	.082	.176
386	.057	.068	.121	.038	.040	.050	.045	.051	.070
1536	.056	.062	.103	.037	.037	.037	.043	.042	.045
6144	.056	.061	.085	.037	.037	.035	.041	.042	.041
12000	.058	.058	.073	.037	.036	.036	.042	.040	.041
24000	.057	.058	.067	.037	.037	.035	.042	.042	.040
100000	.055	.057	.064	.037	.037	.034	.043	.043	.042
Power									
96	.804	.753	.650	.590	.579	.511	.770	.726	.623
386	.784	.731	.630	.583	.566	.489	.721	.686	.583
1536	.771	.716	.615	.581	.567	.482	.671	.642	.548
6144	.762	.711	.604	.583	.566	.485	.639	.620	.531
12000	.751	.704	.595	.582	.564	.485	.637	.615	.528
24000	.746	.699	.593	.584	.565	.484	.637	.613	.529
100000	.748	.694	.592	.588	.565	.484	.639	.612	.528
	Outliers Removed								
Size									
96	.061	.090	.195	.037	.036	.035	.044	.043	.042
386	.057	.060	.095	.036	.036	.035	.043	.041	.043
1536	.054	.053	.057	.035	.038	.033	.040	.044	.041
6144	.054	.053	.056	.040	.037	.037	.044	.044	.043
12000	.052	.053	.054	.038	.035	.033	.041	.042	.041
24000	.052	.052	.053	.036	.039	.034	.041	.045	.041
100000	.052	.052	.053	.037	.035	.034	.042	.042	.042
Power									
96	.906	.931	.927	.706	.713	.656	.772	.776	.719
386	.873	.885	.870	.698	.713	.656	.771	.769	.726
1536	.856	.857	.834	.700	.716	.660	.771	.774	.727
6144	.849	.850	.829	.703	.716	.666	.771	.774	.729
12000	.843	.843	.818	.703	.713	.663	.769	.767	.726
24000	.840	.840	.817	.701	.715	.667	.771	.775	.726
100000	.835	.834	.813	.700	.719	.669	.772	.776	.728

Columns depict the results as F_{ST} varies (.01, .03, and .1) in the candidate markers. The simulated data are a gradient with nine subpopulations; controls are drawn from 1–9 and cases are only from 6–9.

the basis of residence rather than known German ancestry. We removed one case individual who had very different ancestry from the other 415. For 415 cases and 2112 controls, $D = 22$ dimensions were required to explain the significant axes of genetic variation. Many of these axes exhibited extreme outliers (Figure 1A). After removing 53 controls, only three important axes of variation remained. On the basis of the first two eigenvectors, a cluster of cases that differs in ancestry from the control sample was clearly evident (Figure 3A). To identify unmatchable individuals more completely, we computed the distance between each case and the nearest control and vice versa on the basis of three axes of the EVD map. The resulting distribution of distances indicated that 21 cases could not be matched to a control with similar ancestry (Figure 4). By repeating this process of finding the significant eigenvalues and the corresponding minimum distances between cases and controls in the corresponding axes, we subsequently removed an additional one case and 15 controls. After excluding these outliers, only two significant eigenvalues remain when a significance level of 0.01 was used.

Next, with cluster analysis to identify homogeneous strata, 2136 individuals were clustered into 26 strata, each with 20 or more elements and no significant structure within cluster ($p > 0.001$). The remaining 301 individuals were clustered into 24 small clusters. On the basis of these strata, the data were rescaled and the distance between cases and matched controls was determined. Those that were considered unmatchable individuals on the basis of the simulation results were removed (see Appendix). With this process, an additional 20 cases and 48 controls are removed from the dataset for fMatch. The resulting distance between the remaining cases and controls in fMatch is consistent with expectations for cases and controls matched within homogeneous strata (data not shown). In the reduced fMatch sample, two principal axes separate the German control samples by region and define a space, spanned by both cases and controls, that facilitates matching (Figure 3B). These dimensions presumably map onto genetic gradients on the European continent; e.g., the horizontal axis is likely to be related to a north-south gradient

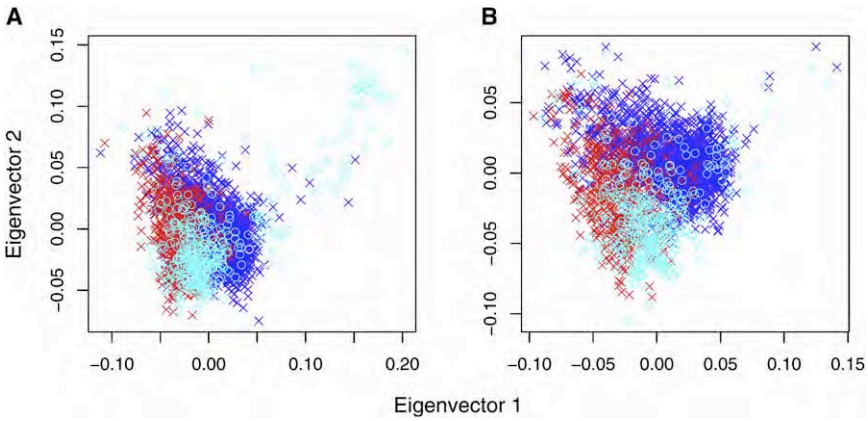

Figure 3. Plots of the First Two Eigenvector Axes for T1D Data before and after Removing the Unmatchable Individuals after Clustering and Rescaling of the Data

Each case (light blue) has a matched control (dark blue = South Germany, red = North Germany) in a close neighborhood after removal of unmatchable individuals; compare before (A) with after (B).

because it tends to separate the German samples by north (PopGen)[12] and south (KORA)[11] origin.[21,22] In the pMatch sample, one additional axis is needed to explain important variation (data not shown).

After final removal of outliers and unmatchable individuals for fMatch, cases and controls were stratified on the basis of their genetic ancestry into 298 strata. Most of the strata (159) contain a single case matched to several controls. A single case matched to a single control occurred in 111 strata. A minority of strata (28) contain a single control matched to multiple cases. For example, in the most extreme strata, a single case was matched to 71 controls and a single control was matched to 13 cases. When a single case is matched to a large number of controls (or vice versa), the information gain from the strata is essentially equivalent to that obtained from a single case matched to a moderate number of individuals. Nevertheless, conditional logistic regression is valid regardless of the lack of balance in the strata. In all, 373 cases were contrasted with 1996 controls by conditional logistic regression (Figure 5, top panel). The results highlight the HLA region, which contains numerous SNPs achieving GWA significance. Variation in the HLA region is well known to account for a large fraction of the risk for T1D.[23–26] No

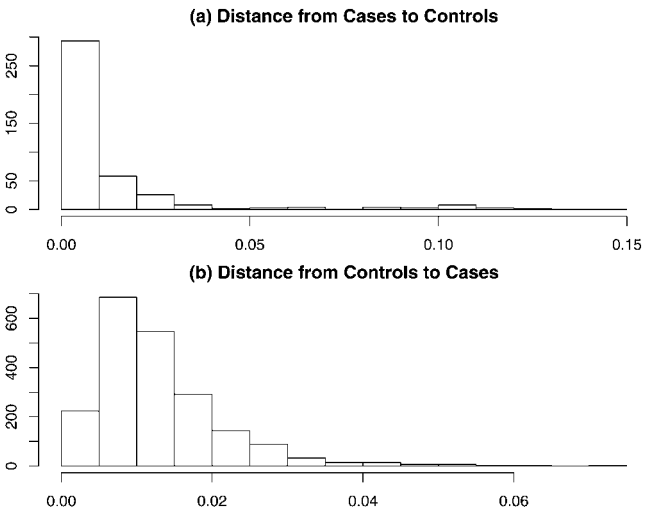

Figure 4. The Distance between Each Case and the Nearest Control and Vice Versa Based on Three Principal Components Are Computed

The distributions differ, and we eliminate 34 cases with distances to the nearest control greater than 0.075. (A) shows the histogram of distances between each case and the nearest control. (B) shows the histogram of distances between each control and the nearest case.

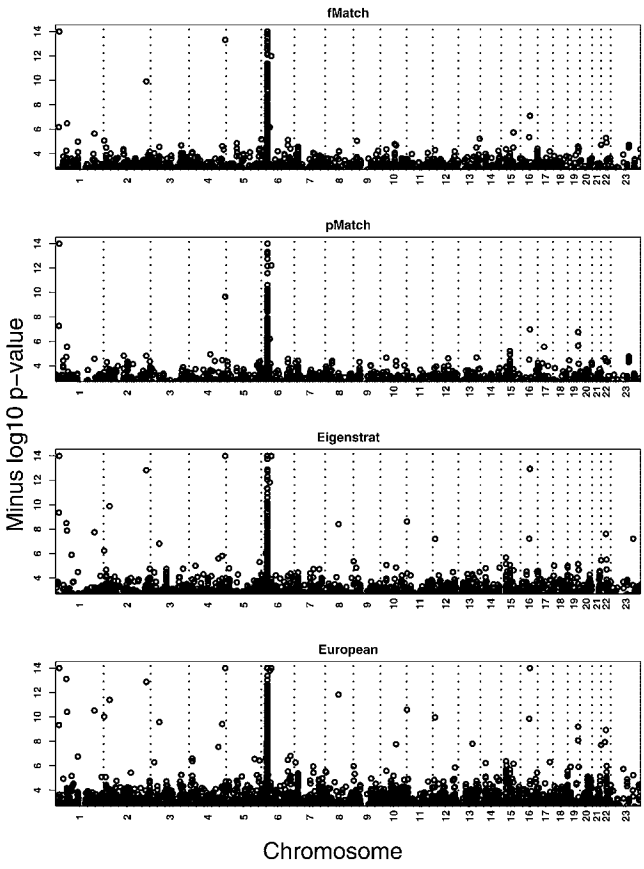

Figure 5. Transformed p Values after Conditional Logit Regression Was Performed on the Data Stratified with fMatch

Transformed p values (negative of the log, base 10). Results from conditional logistic regression on the data stratified with fMatch (top panel) and pmatch (second panel) are shown. Results obtained with Eigenstrat are shown in the third panel. Results obtained when removing observations with very divergent ancestries (inferred with the Eigenstrat rule for outliers) from the bulk of the sample, which was European, are shown in the bottom panel.

other location in the genome contains SNPs with test statistics meeting reasonable criteria for GWA significance ($\leq 10^{-7}$) after ensuring quality genotype calls by visual inspection of the genotype clusters (see Figures S1 and S2 available online for examples). It should be noted, however, that visual inspection of genotype clusters is essential to interpret this Affymetrix "first-generation" genotype data, a feature other GWA studies with this genotyping platform also report.[27]

Results from fMatch agree with our expectations. Other GWA studies have established that all genetic variation thus far uncovered, aside from variation in the HLA region, account for a modest portion of the risk for T1D.[28,29] For detecting loci of modest effect with good power, either sample sizes must be substantial (i.e., thousands of cases and controls genotyped) or a staged study design must be employed. The staged design typically sets a significance level between 0.01 and 0.001 in stage 1, then genotypes all loci meeting this significance level (and quality-control criteria) in a second, larger sample.[30,31] Treating our study as stage 1 with a significance level of 0.007,[30] results from fMatch would include SNPs for genotyping in stage 2 from six out of ten loci now believed to confer risk to T1D.[29] Of the remaining four loci, only one had more than a few SNPs in the region.

Aside from the HLA region, SNPs in or near *PTPN22* (MIM 600716), *IL2RA* (CD25 [MIM 147730]), and *CTLA4* ([MIM 123890]; window = gene location ± 40 Kb) showed enough signal to be passed to stage 2. The smallest p value for each gene was 0.000706 (rs2488457), 0.000995 (rs10905669), and 0.000336 (rs231726). The smallest p values for SNPs close to "risk SNPs" rs2292239 and rs12708716 were 0.00667 (rs2292239) and 0.00539 (rs11647011) for window = SNP location ± 50 Kb. Could it be that the signals in these regions occurred by chance? To answer this question, we performed a simulation experiment. We randomly select from the genome ten intervals that correspond to the same size as the original ten windows (for the HLA region, we assumed a window of 3 Mb). Then, we count the number of intervals in which one or more SNPs have $p < 0.007$ and would thus be genotyped in stage 2. We perform this random selection 10^6 times, counting how many times six or more intervals would have SNPs genotyped in Stage 2. By this experiment, we determined that our results would rarely occur by chance, roughly one in ten thousand times.

A few other observations from these analyses are worth noting. Within the HLA region, Todd[29] cites rs3129934 as the replicated SNP; our independent data and analyses yield a p value for association of 7.2×10^{-10} with this SNP; for the replicated SNP identified in *CTLA4*, rs3087243, our data and analyses yield a p value for association of 0.013, and, as noted above, the replicated SNP rs2292239 produced a p value of 0.00667 from our data. Although the HLA region needs no more support, our results provide further evidence for replication in *CTLA4* and at rs2292239. For genotype cluster plots for the cited SNPs, see Figure S1. In addition, for all of the loci cited above, we have compared our data to that reported by the Wellcome Trust Case Control Consortium.[27] For these loci, the allele in excess in cases is the same for both data sets (data not shown).

Four loci did not pass stage 1 criteria. None of these SNPs reported by Todd et al.[29] as risk loci were on our Affymetrix genotyping array. Of these four risk SNPs, only rs1893217 in 18p11 was covered well in terms of genotyped SNPs in substantial linkage disequilibrium (LD) with it. This SNP is in almost complete LD with rs2542151 according to HapMap; it passed our QC, but it shows no evidence for association in our data ($p = 0.51$). For the proinsulin precursor gene, *INS* (MIM 176730), only two SNPs on the array pass QC and fall in the region, but HapMap contains no information about their LD with the reported risk SNP, rs689, and they show no association ($p > 0.35$). For the gene encoding interferon-induced helicase C domain-containing protein 1, *IFIH1* (MIM 606951), the reported risk allele shows modest LD with a SNP we genotyped, namely rs7608315, which shows no association ($p = 0.38$). Finally, for the 12q24 region, rs3184504 is identified as the risk SNP. One SNP in this region passed QC for our data, and it is modestly associated with risk for T1D ($p = 0.046$).

The vast majority of the SNPs from this or any relevant GWA are independent of risk for T1D. Many SNPs from the HLA regions of chromosome 6 are associated, however. After eliminating HLA SNPs, ~5% of the association tests are expected to have p values <0.05. Of the 284,216 tests, 7.0% were significant at $\alpha = 0.05$ for fMatch. A moderate excess of false positives occurs for any reasonable choice of α. Given the success of the GEM method in the simulations, in terms of controlling the false-positive rate, we wondered whether the source of additional false positives could be poor-quality genotype calls. Indeed, by assessing genotype clusters for all SNPs producing p values $\leq 10^{-4}$, we find a rate of poor calls of 60%–67% (Figure S2). The rate of poor-quality genotype calls increases as the p value decreases. Predominantly, the problematic calls occur for the T1D sample. On the basis of our estimated rate of poor-quality genotype calls, we believe the excess false-positive rate is attributable to data-quality issues, not the method.

We also analyzed GWA data by using pMatch and Eigenstrat and by ignoring population substructure after discarding outliers with the Eigenstrat rule (see Figure 5). As expected, pMatch shows the lowest rates of positive findings, whereas ignoring structure yields the most. Like fMatch, it appears the excess of false positives for pMatch is due to poor-quality genotype calls. The same is predominantly true for Eigenstrat, but we note that an ample number of SNPs producing small p values are not attributable to poor quality, and this problem is amplified by ignoring structure. At significance level 0.0001, after visual inspection of genotypes fMatch has half the false-positive rate of Eigenstrat.

To further validate GEM, we tried a null experiment. We randomly labeled half of the KORA data as cases and half as controls and repeated the matching analysis. Removal of 72 outliers reduced the number of significant eigenvalues

required to explain the variation from 24 to 2. After this simplification, only 12 unmatchable individuals remained. All three methods of analysis (Eigenstrat, Pmatch, and fMatch) produced type I error rates that were on target.

Discussion

Our GWA analyses of T1D are meant to accomplish two goals. First, they illustrate the utility of ancestry matching in the face of a very difficult problem, that being when cases are sampled in a region quite different from the region of the controls. In our case, the T1D sample comes from any American of nominal European ancestry, whereas the controls were recruited among residents of Germany. Such constellations can also arise even if cases and controls are sampled from the same geographical region. We would expect the example to be especially salient for American samples. Second, we wished to use the results to evaluate reported T1D risk loci and, in later analyses, discover new loci. The results show that genetic or ancestry matching can be an important ingredient in the toolbox of researchers who are performing GWA analyses. Moreover, our results do lend support for previous GWA findings for T1D.[28,29]

We do not yet know whether our analyses have identified any new risk loci for T1D. Although it seems unlikely given the modest sample of cases, a substantial number of controls have been analyzed. Moreover, for a rare disease like T1D, using unscreened instead of screened controls has almost no impact on power.[32] We plan various kinds of stage 2 analyses to assess the association signals from our GWA results. In addition, by agreement the data generated by our project will be reported back to the GoKinD database, and GoKinD will make the data available to qualified investigators. Thus, these data will shortly be available to the research community, and we will be pleased to share detailed results upon request.

We have described how to use genetic matching to enhance a case-control study. We note, however, that these methods can also be used for the analysis of quantitative traits. Once homogeneous clusters are identified, they can be entered into a model as block effects, and the quantitative trait can be analyzed with standard statistical tools, such as analysis of variance.

Theory, simulations, and real-data analyses suggest that genetic matching is useful and powerful for GWA, especially when the samples of cases and controls cannot be guaranteed to be drawn from the same population. It can diminish the false-positive rate, sometimes substantially, and have only modest impact on power. Among others,[33–36] methods similar to Eigenstrat[4] also limit the impact of population structure, but for challenging designs, they cannot be expected to completely control the false-positive rate. Perhaps the gold standard for GWA studies should be to evaluate the data with both regression methods such as Eigenstrat and epidemiological methods such as fMatch. When the results of these methods agree, researchers have greater assurance of validity; it is when the results diverge that we should be wary.

Appendix

EVD of Allele Counts

Using allele counts for SNPs $l = 1, \ldots, L$, and individuals $i = 1, \ldots, N$, create an $N \times L$ matrix X. For p_l, the lth allele frequency, center allele counts in column l by subtracting $2p_l$ and scale by dividing by $(2p_l(1-p_l))^{1/2}$. Find the EVD of $XX^t = U\lambda U^t$. In the D dimensional space defined by the top D eigenvectors, the "ancestry" value for the ith subject is determined by the ith row of the eigenvectors $u_{i1}, \ldots, u_{iD}$. The dth eigenvalue, λ_d, determines the scaling of distances in the dth dimension. These coordinates are used for matching.

Model for Population Stratification

The mean of allele frequencies from a set of populations is assumed to be the allele frequency of an ancestral population. Individual populations have each diverged from the ancestral population over time, with fixation index F_{ST}, a measure of population differentiation. Within a subpopulation j, suppose that allele counts are independent and identically distributed and that allele a is drawn with probability p_j. If X is counting allele a, then $X \sim \text{Binomial}(2,p_j)$. Let P be the random variable that varies across subpopulations, with p_j as the realized value in subpopulation j: $P \sim Beta(\alpha_1,\alpha_2)$, $\alpha_1 + \alpha_2 = 1/F_{ST} - 1$. Assume that we have the minor allele frequencies of an ancestral population *p.loci* (in our simulations *p.loci* is uniform between .05 and .5) at L loci. From the ancestral population J, subpopulations have been formed. By knowing F_{ST}, for each marker l we can define $\alpha_{1,l} = p.loci_l \times (1/F_{ST} - 1)$ and $\alpha_{2,l} = (1 - p.loci_l) \times (1/F_{ST} - 1)$ and generate the alleles as described above. When used in simulation studies, this is often called the Balding-Nichols model.[17] For simulation of a cline (or a gradient), it is enough to order p_{jl} so that $p_{1l} \leq \ldots \leq p_{Jl}$ for each l.

Hypothesis Test for Population Structure

A formal significance test for population structure is based on a theoretical result for the eigenvalue distribution of a null sample covariance matrix.[6,37] For a homogeneous population, the largest eigenvalue, properly normed, approximately follows the Tracy-Widom distribution[37] $W_d = (\lambda_d - \mu_{NL})/\sigma_{NL}$ with centering and scaling parameters that depend on both N and L, $\mu_{NL} = ((L-1)^{1/2} + N^{1/2})^2$ and $\sigma_{NL} = ((L-1)^{1/2} + N^{1/2})(1/((L-1)^{1/2}) + 1/N^{1/2})^{1/3}$. We can test the null hypothesis of population homogeneity against an alternative hypothesis of population heterogeneity. The sample covariance matrix S follows a $(N-1) \times (N-1)$ Wishart distribution. The test for population structure will be applied iteratively (i.e., the leading eigenvalue, then the second and so on). If we find the first d eigenvalues $\lambda_1, \ldots, \lambda_d$ to be significant, we test λ_{d+1} as

though S were an $(N-d-1)\times(N-d-1)$ Wishart matrix. If an eigenvalue is not significant, the smaller eigenvalues will not be significant either.

Removing Unmatchable Individuals

EVD determines the distance between individuals on the basis of the top D eigenvectors and eigenvalues. To stabilize the distance metric, we use the normed eigenvalues, W_d, plus a constant a, chosen to ensure the weights are positive. The distance between individuals i and i' is calculated as $g(i,i') = \left\{\sum_{d=1}^{D}(W_d + a)(u_{id} - u_{i'd})^2\right\}^{1/2}$.

To rescale the distances, let $S_k \subset \{1,2,\ldots,N\}$ be the indices of individuals in the k'th cluster. Let r_k be the number of individuals in the k'th cluster. For scaling subject $i \in S_k$, we use the eigenvector values $(u_{i1},\ldots,u_{iD})$ but not the eigenvalues. Assume that the eigenvector representation of each individual consists of an ancestry signal plus random noise: $u_{id} = \mu_{id} + \varepsilon_{id}$.

For homogeneous data, because all individuals came from a common source, the ancestry signal is 0 and the representation consists simply of random noise $u_{id} = \varepsilon_{id}$. Our target is to identify approximately homogeneous subpopulations that have little or no diversity for ancestry. If the clustering is successful, the signal of each individual in subset S_k can be approximated by $\overline{u}_{dk} = \sum_{i\subset S_k} u_{id}/r_k$, and the noise can be approximated by $u_{id} - \overline{u}_{dk}$. But notice that EVD automatically scales the eigenvectors so that $\sum_i u_{id}^2 = 1$ and $\overline{u}_d = 0$. A traditional sum of squares decomposition leads to

$$1 = \sum_i u_{id}^2 = \sum_k \sum_{i\in S_k} (u_{id} - \overline{u}_{dk})^2 + \sum_k r_k \overline{u}_{dk}^2,$$

i.e., the total sum of squares (SSTotal) equals the sum of squares attributable to random variation or error (SSError) plus the sum of squares attributable to ancestry differences (SSModel). Unit scaling of SSTotal causes the distances between individuals from heterogeneous populations to be uncomparable to distances in homogeneous populations. For example, if the sample derives from two highly differentiated populations so that SSError = 0.01 and SSModel = 0.99, then the expected distance between two individuals with common ancestry is ~0.01/n. Alternatively, if the populations have identical ancestry, then the expected distance between two individuals is ~1/n. For comparing to a homogeneous scaling, we wish to rescale the random noise so that SSError is 1. It follows that the data will be rescaled equivalently to homogeneous data if we set $c_d^2 = \sum_k \sum_{i\in S_k} (u_{id} - \overline{u}_{dk})^2$ and rescale the data such that $u_{id}^* = u_{id}/c_d$.

In practice, $\overline{u}_{dk}$ provides a good estimate of the signal only when the cluster size is sufficiently large, say greater than 10. Hence, to compute c_d^2, include only those clusters S_k including 10 or more elements in the sum and then multiply by $n/\sum_k (r_k - 1)$ to account for the missing clusters. Notice that we scale differently for each of the d dimensions to stretch and shrink accordingly to get the proper scaling of the data.

In the final step, find the distances between individuals with the u_{id}^* instead of u_{id} and use the expected value of normed eigenvalues $W_1,\ldots,W_D$ obtained from the simulation, instead of the actual eigenvalues. Match rescaled data with fMatch or pMatch and measure the distances between cases and controls. Any individuals with distances in this metric exceeding the 99.9th quartile of the null distribution of distances are declared unmatchable.

Supplemental Data

Two figures are available at http://www.ajhg.org/.

Acknowledgments

The Genetics of Kidneys in Diabetes (GoKinD) Study sample collection is supported by the Juvenile Diabetes Research Foundation in collaboration with the Joslin Diabetes Center and George Washington University and by the United States Centers for Disease Control and Prevention. This work was funded by the National Institutes of Health grant MH057881 awarded to B.D. and K.R. and by the Department of Defense (grant W81XWH-07-1-0619) awarded to M.T. The MONICA/KORA Augsburg studies were financed by the GSF-National Research Center for Environment and Health, Neuherberg, Germany, and supported by grants from the German Federal Ministry of Education and Research (BMBF). Part of this work was financed by the German National Genome Research Network (NGFN). The genotypes reported in this publication were generated independently from any other initiatives supported by other organizations.

Received: October 1, 2007
Revised: November 16, 2007
Accepted: November 20, 2007
Published online: January 24, 2008

Web Resources

The URLs for data presented herein are as follows:

CHIAMO, http://www.stats.ox.ac.uk/%7Emarchini/software/gwas/chiamo.html
dbGaP, http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap
GEM, http://wpicr.wpic.pitt.edu/WPICCompGen/
GoKinD, http://www.jdrf.org/gokind
HapMap Frequencies, http://www.hapmap.org/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim
Optmatch, http://cran.r-project.org/doc/packages/
T1Dbase, http://t1dbase.org/

References

1. Lee, W.-C. (2004). Case-control association studies with matching and genomic controlling. Genet. Epidemiol. *27*, 1–13.
2. Hinds, D.A., Stokowski, R.P., Patil, N., Konvicka, K., Kershenobich, D., Cox, D.R., and Ballinger, D.G. (2004). Matching strategies for genetic association studies in structured populations. Am. J. Hum. Genet. *74*, 317–325.

3. Zhang, S., Zhu, X., and Zhao, H. (2003). On a semiparametric test to detect associations between quantitative traits and candidate genes using unrelated individuals. Genet. Epidemiol. *24*, 44–56.
4. Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. *38*, 904–909.
5. Rosenbaum, P.R. (1995). Observational Studies (New York: Springer-Verlag).
6. Patterson, N., Price, A.L., and Reich, D. (2006). Population structure and eigenanalysis. PLoS Genet. *2* 10.1371/journal.pgen.0020190.
7. Jolliffe, I.T. (2002). Principal Component Analysis (New York: Springer).
8. Venables, W.N., and Ripley, B.D. (2002). Modern Applied Statistics with S, Fourth Edition (New York: Springer).
9. Everitt, B.S. (1993). Cluster Analysis (London: Edward Arnold).
10. Mueller, P.W., Rogus, J.J., Cleary, P.A., Zhao, Y., Smiles, A.M., Steffes, M.W., Bucksa, J., Gibson, T.B., Cordovado, S.K., Krolewski, A.S., et al. (2006). Genetics of Kidneys in Diabetes collection available for identifying genetic susceptibility factors for diabetic nephropathy in type 1 diabetes. J. Am. Soc. Nephrol. *17*, 1782–1790.
11. Wichmann, H.E., Gieger, C., Illig, T., and MONICA/KORA Study Group. (2006). KORA-gen–resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen *67* (*Suppl 1*), S26–S30.
12. Krawczak, M., Nikolaus, S., von Eberstein, H., Croucher, P.J., El Mokhtari, N.E., and Schreiber, S. (2006). PopGen: Population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. Community Genet. *9*, 55–61.
13. Steffens, M., Lamina, C., Illig, T., Bettecken, T., Vogler, R., Entz, P., Suk, E.K., Toliat, M.R., Klopp, N., Caliebe, A., et al. (2006). SNP-based analysis of genetic substructure in the German population. Hum. Hered. *62*, 20–29.
14. Ringquist, S., Styche, A., Rudert, W.A., and Trucco, M. (2007). Pyrosequencing-based strategies for improved allele typing of human leukocyte antigen loci. Methods Mol. Biol. *373*, 115–134.
15. Rabbee, N., and Speed, T.P. (2006). A genotype calling algorithm for affymetrix SNP arrays. Bioinformatics *22*, 7–12.
16. Marchini, J., Howie, B., Myers, S., McVean, G., and Donnelly, P. (2007). A new multipoint method for genome-wide association studies via imputation of genotypes. Nat. Genet. *39*, 906–913.
17. Balding, D.J., and Nichols, R.A. (1995). A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity. Genetica *96*, 3–12.
18. Pasquali, L., Fan, Y., Trucco, M., and Ringquist, S. (2006). Rehabilitation of adaptive immunity and regeneration of beta cells. Trends Biotechnol. *24*, 516–522.
19. Lowel, H., Doring, A., Schneider, A., Heier, M., Thorand, B., and Meisinger, C. (2005). The MONICA Augsburg surveys–basis for prospective cohort studies. Gesundheitswesen *67* (*Suppl 1*), S13–S18.
20. Rinaldo, A., Bacanu, S.-A., Devlin, B., Sonpar, V., Wasserman, L., and Roeder, K. (2005). Characterization of multilocus linkage disequilibrium. Genet. Epidemiol. *28*, 193–206.
21. Cavalli-Sforza, L.L., Menozzi, P., and Piazza, A. (1994). The History and Geography of Human Genes (Princeton, NJ: Princeton University Press).
22. Rosenberg, N.A., Mahajan, S., Ramachandran, S., Zhao, C., Pritchard, J.K., and Feldman, M.W. (2005). Clines, clusters and effect of study design on the inference of human population structure. PLoS Genetics *1*, e70.
23. Todd, J.A., Bell, J.I., and McDevitt, H.O. (1987). HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature *329*, 599–604.
24. Morel, P.A., Dorman, J.S., Todd, J.A., McDevitt, H.O., and Trucco, M. (1988). Aspartic acid at position 57 of the HLA-DQ beta chain protects against type I diabetes: A family study. Proc. Natl. Acad. Sci. USA *85*, 8111–8115.
25. Dorman, J.S., LaPorte, R.E., Stone, R.A., and Trucco, M. (1990). Worldwide differences in the incidence of type I diabetes are associated with amino acid variation at position 57 of the HLA-DQ beta chain. Proc. Natl. Acad. Sci. USA *87*, 7370–7374.
26. Hyttinen, V., Kaprio, J., Kinnunen, L., Koskenvuo, M., and Tuomilehto, J. (2003). Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs: A nationwide follow-up study. Diabetes *52*, 1052–1055.
27. Wellcome Trust Case Control Consortium (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature *447*, 661–678.
28. Hakonarson, H., Grant, S.F., Bradfield, J.P., Marchand, L., Kim, C.E., Glessner, J.T., Grabs, R., Casalunovo, T., Taback, S.P., Frackelton, E.C., et al. (2007). A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. Nature *448*, 591–594.
29. Todd, J.A., Walker, N.M., Cooper, J.D., Smyth, D.J., Downes, K., Plagnol, V., Bailey, R., Nejentsev, S., Field, S.F., and Payne, F. (2007). Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat. Genet. *39*, 857–864.
30. Wang, H., Thomas, D.C., Pe'er, I., and Stram, D.O. (2006). Optimal two-stage genotyping designs for genome-wide association scans. Genet. Epidemiol. *30*, 356–368.
31. Skol, A.D., Scott, L.J., Abecasis, G.R., and Boehnke, M. (2007). Optimal designs for two-stage genome-wide association studies. Genet. Epidemiol. *31*, 776–788.
32. Bacanu, S.A., Devlin, B., and Roeder, K. (2000). The power of genomic control. Am. J. Hum. Genet. *66*, 1933–1944.
33. Devlin, B., and Roeder, K. (1999). Genomic control for association studies. Biometrics *55*, 997–1004.
34. Pritchard, J., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics *155*, 945–959.
35. Devlin, B., Bacanu, S.A., and Roeder, K. (2004). Genomic control to the extreme. Nat. Genet. *36*, 1129–1130.
36. Epstein, M.P., Allen, A.S., and Satten, G.A. (2007). A simple and improved correction for population stratification in case-control studies. Am. J. Hum. Genet. *80*, 921–930.
37. Johnstone, I. (2001). On the distribution of the largest eigenvalue in principal components analysis. Annals of Statistics *29*, 295–327.